(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
9 January 2003 (09.01.2003)

PCT

(10) International Publication Number
WO 03/003012 A1

(51) International Patent Classification[7]: G01N 33/53, C12Q 1/02, C07C 59/64

(21) International Application Number: PCT/AU02/00856

(22) International Filing Date: 28 June 2002 (28.06.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PR 5986 29 June 2001 (29.06.2001) AU

(71) Applicant *(for all designated States except US)*: **MEDI-MOLECULAR PTY. LTD.** [AU/AU]; Science Park Adelaide, Bedford Park, South Australia 5042 (AU).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **JAMES, Robert** [AU/AU]; 23A Carunta Street, Wattle Park, South Australia 5066 (AU). **EDDIE, Lawrence** [AU/AU]; 35 Highland Drive, Bellevue Heights, South Australia 5050 (AU). **KAZENWADEL, Jan** [AU/AU]; 4 Thorngate Drive, Belair, South Australia 5052 (AU). **O'CONNOR, Susan** [AU/AU]; 3 Clifton Street, Prospect, South Australia 5082 (AU). **RAZZINO, Pasquale** [AU/AU]; 191 Greenhill Road, Parkside, South Australia 5063 (AU). **WARD, David** [AU/AU]; 2/16 Conyngham Street, Glenside, South Australia 5065 (AU).

(74) Agent: **PHILLIPS ORMONDE & FITZPATRICK**; 367 Collins Street, Melbourne, VIC 3000 (AU).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/003012 A1

(54) Title: IDENTIFICATION OF INTERACTING MOLECULES

(57) Abstract: The present invention provides a method for identifying a protein capable of binding to a specific target molecule. The method involves allowing candidate proteins to bind to the target molecule in the presence of a second molecule which is structurally similar to the non-nucleic acid target molecule, but deficient in a desired activity of the target molecule, and isolating the proteins that bind to the target molecule. The invention also provides analogues of flurbiprofen and sulindac as target molecules for use in the methods of the invention.

IDENTIFICATION OF INTERACTING MOLECULES

Field of the Invention

The present invention relates to methods for identifying candidate proteins that interact with target substrates. The present invention also relates to candidate proteins identified using such methods, and to target molecules that can be used to identify candidate proteins.

Background of the Invention

The identification of the interactions between molecules is central to the understanding of how biologically active molecules exert their effect. In most cases, the molecules to which a biologically active molecule binds are not known. For this reason, the identification of the interacting partner molecule is an important step in the understanding of how a particular molecule exerts its effect. For example, many drugs have been identified for which the mechanism of action has not been determined. The identification of the molecule (or molecules) with which the drug interacts may allow further understanding of how the drug works, and therefore lead to the design or discovery of new drugs with enhanced efficacy.

The identification of molecules that interact with a target may be achieved in a number of ways. For example, *in vitro* methods such as affinity purification may be used. Alternatively, *in vivo* interaction assays may be used to identify interacting molecules. When the interacting molecules are proteins, genetic analyses may also reveal the identity of interacting proteins.

In vitro methods generally rely on the strength (affinity) of the interaction between the molecules to separate the binding molecule from a complex pool of non-binding molecules. For example, affinity purification methods generally rely on the principal of immobilisation of a target molecule on a solid support, passing a mixture of molecules over the solid support, washing the solid support so as to

remove molecules that do not bind, and isolating the molecules that remain bound to the immobilised target. Molecules that remain bound to the target are generally ones that bind to the target molecule with a higher affinity, and are therefore more likely to be of biological relevance. The lower affinity binding molecules that are removed by the selected washing treatment are usually less likely to be of biological relevance.

A variety of molecules may be used as targets in affinity purification systems. Such targets include small organic molecules, cofactors, nucleic acids, small peptides, proteins, oligosaccharides and lipids.

Affinity purification methods may generally be used to isolate any type of molecule that interacts with a target molecule of interest. As many biologically active molecules exert their effects by binding to proteins, affinity purification methods have been used to isolate proteins that bind to a specific target. In such cases, the candidate binding protein may often need to be isolated from a complex mixture of proteins.

Methods have become available whereby very complex pools of candidate binding molecules may be generated for the purposes of screening. This may be important in cases where a candidate binding molecule is likely to be present at a very low abundance. In such a case, the likelihood of finding an appropriate binding molecule will increase with increasing complexity of the mixture to be screened. Methods for generating complex mixtures of potential binding molecules include, for example, the chemical synthesis of short random peptides and the chemical synthesis of short random nucleic acid molecules (aptamers).

Complex mixtures of proteins may be generated by the cloning of a large number of individual cellular DNAs into expression systems to form a library of proteins expressed from the cloned DNA molecules. For example, a pool of complementary DNAs (cDNAs) may be isolated from mRNAs isolated from a particular source, the cDNAs cloned into an expression vector and the proteins encoded by each of the cDNAs expressed. The proteins expressed from such

pools or libraries may be used as the source of potential binding proteins for screening by affinity purification.

One method for screening complex mixtures of proteins using an affinity purification approach is the method referred to as "phage display". In this method, various DNAs are cloned in a viral nucleic vector. The DNAs are inserted into a viral gene that normally expresses a protein that is found on the surface of the virus. When viral particles are produced, the protein expressed from the particular DNA inserted into the virus DNA will be displayed on the surface of the virus. If a library of DNA molecules is cloned into the viral vector, each virus produced will display a different protein from the library on its surface.

The complex mixture of viral particles so produced may be passed over a target molecule immobilised to a solid support. To identify candidate proteins that may bind to the target, the solid support is washed and the viral particles that remain bound are collected. Phage display methods are useful for the screening of potential candidate binding proteins, because once the viral particles binding to the target are isolated, the viral particles may be allowed to infect new cells and produce a new enriched population of viral particles. The resultant new mixture of viral particles may then be re-applied to the immobilised target molecule and the process reiterated. In this way, a population of viral particles may be successively enriched for those viral particles expressing candidate binding proteins on their surface.

However, a deficiency with the use of affinity purification methods such as phage display has been the inability to readily distinguish between the binding of *bona fide* candidate molecules and the binding of other molecules that are not biologically relevant. This may occur for a number of reasons. Many proteins may bind to the target molecule with high affinity, but the binding may be for reasons that are unrelated to the biological activity of the target molecule. Alternatively, the washing regime selected may not be sufficiently capable of discriminating between the binding of *bona fide* candidate proteins and the binding of other proteins with reduced affinity for the target. Additionally, some

interactions with other proteins may also interfere with the binding of *bona fide* candidate proteins.

The present invention relates to an improved method for isolating proteins from complex mixtures, by providing means to improve the likelihood that the proteins identified are biologically relevant to the desired activity or function of the target molecule.

Summary of the Invention

The present invention provides a method for identifying a protein capable of binding to a target molecule, the method including the steps of:

(a) providing a pool of candidate proteins;

(b) providing a non-nucleic acid target molecule, wherein the non-nucleic acid target molecule is coupled to a selectable moiety;

(c) providing a second molecule which is structurally similar to the non-nucleic acid target molecule, wherein the second molecule is deficient in a desired activity of the target molecule;

(d) allowing one or more of the candidate proteins to bind the non-nucleic acid target molecule in the presence of the second molecule;

(e) isolating a protein bound to the target molecule; and

(f) identifying the binding protein.

The present invention provides a method for identifying a protein capable of binding to a target molecule, the method including the steps of:

(a) providing a pool of candidate proteins, wherein each candidate protein is displayed on the surface of a viral particle;

(b) providing a non-nucleic acid target molecule, wherein the non-nucleic acid target molecule is coupled to a selectable moiety;

(c) providing a second molecule which is structurally similar to the non-nucleic acid target molecule, wherein the second molecule is deficient in a desired activity of the target molecule;

(d) allowing one or more of the candidate proteins to bind to the non-nucleic acid target molecule in the presence of the second molecule;

(e) isolating one or more proteins bound to the target molecule;

(f) amplifying the viral particles encoding the isolated binding proteins;

(g) reiterating steps (a) through (f); and

(h) identifying the binding protein.

The present invention provides a method for identifying a protein capable of binding to target molecule, the method including the steps of:

(a) providing a first pool of candidate proteins;

(b) providing a non-nucleic acid target molecule, wherein the non-nucleic acid target molecule is coupled to a selectable moiety;

(c) providing a second molecule which is structurally similar to the non-nucleic acid target molecule, wherein the second molecule is deficient in a desired activity of the target molecule;

(d) allowing one or more of the candidate proteins in the first pool to bind to the non-nucleic acid target molecule in the presence of the second molecule;

(e) isolating a protein in the first pool that binds to the target molecule;

(f) comparing the level of the protein in the first pool of candidate proteins with the level of the protein in a second pool of proteins; and

(g) identifying a protein that is differentially represented between the first and second pools.

The present invention provides a method for identifying a protein capable of binding to a target molecule, the method including the steps of:

(a) providing first and second pools of candidate proteins;

(b) providing a non-nucleic acid target molecule, wherein the non-nucleic acid target molecule is coupled to a selectable moiety;

(c) providing a second molecule which is structurally similar to the non-nucleic acid target molecule, wherein the second molecule is deficient in a desired activity of the target molecule;

(d) allowing one or more of the candidate proteins in the first pool to bind to the non-nucleic acid target molecule in the presence of the second molecule;

(e) isolating one or more proteins in the first pool that bind to the target molecule;

(f) allowing one or more of the candidate proteins in the second pool to bind to the non-nucleic acid target molecule in the presence of the second molecule;

(g) isolating one or more proteins in the second pool that bind to the target molecule; and

(h) comparing one or more proteins isolated from each of the first and second pools to identify a protein that is differentially represented between the first and second pools.

The present invention relates to a method for identifying proteins that are able to bind to a target molecule, the target molecule preferentially being a biologically active molecule. The identification of proteins that are able to bind to such a target molecule may be important for a number of reasons. For example, it may allow the identification of the proteins that the target molecule binds to in order to exert its biological effect. It may also allow the identification of new proteins that themselves may serve as drugs to inhibit or augment the biological activity of the target molecule.

It has been determined by the applicant that when screening a large number of candidate proteins for their ability to bind to a target molecule, the presence in the binding reaction of a second molecule, which is structurally similar to the target molecule but deficient in a biological activity of the target molecule, may allow improved detection of proteins that are involved in a biologically relevant interaction with the target molecule.

This improved effect is particularly apparent when the candidate binding proteins are screened by a method in which the proteins are displayed on the surface of a viral particle. It also appears that reiterated cycles of binding the protein to the

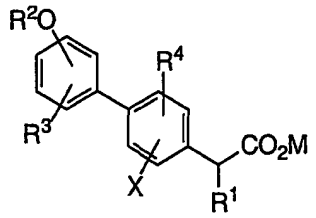
target molecule, the isolation of the viral particles that contain the binding protein and the amplification of these viral particles, may allow an enrichment of proteins that bind to the target in a way that may be biologically important.

The presence of a second molecule in the binding reaction, which is similar to the target molecule but which is deficient in a biological activity of the target molecule, may diminish the binding to the target molecule of proteins that inhibit the binding of biologically relevant proteins. In this manner, the use of the second molecule may serve to diminish a number of inhibitory interactions and thereby may allow binding of biologically relevant proteins. The use of reiterated cycles of binding and amplification of viral particles may further augment these effects of the second molecule.

Thus, the methods according to the present invention allow the identification of candidate proteins that are more likely to interact with a target molecule in a manner that is associated with a desired activity of the target molecule. The methods of the present invention also allow the identification of candidate proteins that interact with a specific region of a target molecule. In addition, the method of the present invention allows the identification of candidate proteins that are not only more likely to be biologically relevant, but which are also differentially represented between pools of candidate proteins.

In a preferred form of the invention the target molecule is either of the non-steroidal antiinflammatory drugs, flurbiprofen or sulindac sulfide or biologically active analogues of either of these. In this form of the invention the second molecule may be an analogue of either flurbiprofen or sulindac sulfide that is inactive or has a reduced biological activity relative to the parent compound.

Accordingly, the present invention provides an analogue of flurbiprofen, the analogue having the formula (I):

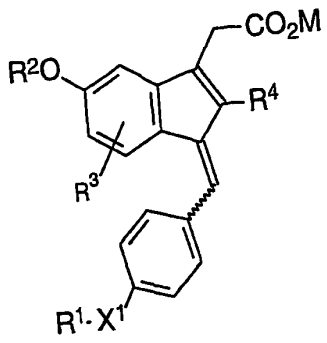


(I)

or a salt thereof, wherein:

- R^1 is selected from hydrogen and lower alkyl (C1 to C8);
- R^2 is $YX^2((CH_2)_m X^2)_n$-, wherein m is 2 to 4, n is 1 to 6, X^2 is selected from O, S and N, and Y is independently selected from hydrogen, lower alkyl, or a suitable heteroatom protecting group;
- R^3 is selected from one or more of hydrogen, alkyl, aryl, halogen, hydroxy, alkoxy, aryloxy, amino (unsubstituted and substituted) and caroboxy;
- R^4 is selected from one or more of hydrogen, alkyl, aryl, halogen, hydroxy, alkoxy, aryloxy, amino (unsubstituted and substituted) and caroboxy;
- X is selected from fluoro, chloro, bromo and iodo;
- M is selected from hydroxy, alkoxy, aryloxy, amino, alkylamino (mono- and di-), arylamino (mono- and di-), N-morpholino, hydroxyalkylamino, dialkylaminoalkylamino, aminoalkylamino, polyhydroxyamino, and salts of any of the aforementioned.

Alternatively or in addition, the present invention provides an analogue of sulindac sulfide, the analogue having the formula (II):



(II)

or salts thereof, wherein:

- X^1 is selected from sulfide, sulfone and sulfoxide;
- R^1 is selected from hydrogen, hydroxy (when X^1 is sulfone or sulfoxide), and lower alkyl (C1 to C8);
- R^2 is $YX^2((CH_2)_m X^2)_n$-, wherein m is 2 to 4, n is 1 to 6, X^2 is selected from O, S and N, and Y is independently selected from hydrogen, lower alkyl, or a suitable heteroatom protecting group;
- R^3 is selected from hydrogen, halogen, alkyl, alkoxy, acyloxy, amino, alkylamino (mono- and di-), arylamino (mono- and di-), nitro, cyano, carboxyl;
- R^4 is selected from hydrogen and lower alkyl (C1 to C8); and
- M is selected from hydroxy, alkoxy, aryloxy, amino, alkylamino (mono- and di-), arylamino (mono- and di-), N-morpholino, hydroxyalkylamino, dialkylaminoalkylamino, aminoalkylamino, polyhydroxyamino, and salts of any of the aforementioned.

The compounds of formula (I) or (II) may be used as target molecules or second molecules to identify proteins that interact with flurbiprofen or sulindac respectively.

Various terms that will be used throughout this specification have meanings that will be well understood by a skilled addressee. However, for ease of reference, some of these terms will now be defined.

The term "non-nucleic acid target molecule" as used throughout the specification is to be understood to mean any molecule to which a protein may bind, but which is not a nucleic acid molecule. For example, the target molecule may include drug molecules, proteins, peptides, polypeptides, polysaccharides, glycoproteins, hormones, receptors, lipids, small molecules, metabolites, cofactors, transition state analogues and toxins.

In terms of the present invention, the second molecule will be a molecule that is structurally similar to the non-nucleic acid target molecule. The term "structurally similar" as used throughout the specification is to be understood to mean a molecule that is similar to the non-nucleic acid target molecule in terms of its three dimensional structure.

For example, structurally similar molecules may include isomeric molecules such as isomers, geometric isomers, enantiomers, conformers, stereoisomers, structural isomers, molecules that substitute one or more chemical groups in a molecule with other chemical groups, or molecules that are substantially similar in the three dimensional structure of one or more parts of the molecule.

The term "protein" as used throughout the specification is to be understood to mean any polypeptide consisting of two or more constituent amino acids. The polypeptide may also contain one or more side chains derived from a modified amino acid.

The term "viral particle" as used throughout the specification is to be understood to mean any virus with a protein coat, wherein the virus genome contains a gene for a coat protein that will allow the display of a subject protein on the surface of the protein, when the DNA encoding the subject protein is inserted into the gene for an appropriate coat protein. For example, the viral particle according to the present invention may include bacteriophage particles.

The term "selectable moiety" as used throughout the specification is to be understood to mean any chemical group that is linked to the target molecule and which may be used to substantially purify the target molecule.

General Description of the Invention

As mentioned above, in one form the present invention provides a method for identifying a protein capable of binding to a target molecule, the protein so identified being more likely to be interact with the target molecule in a manner that is associated with a desired activity of the target molecule.

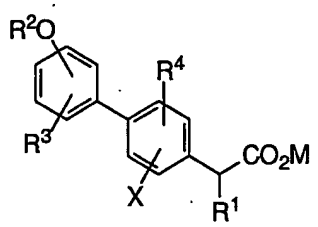
The target molecule according to the methods of the present invention may be any molecule to which a protein may bind, but which is not in itself a nucleic acid molecule. For example, the target molecule may include drug molecules, proteins, peptides, polypeptides, polysaccharides, glycoproteins, hormones, receptors, lipids, small molecules, metabolites, cofactors, transition state analogues and toxins.

The target molecule will have an activity or function that is of interest, referred herein to as "a desired activity" of the target molecule. The activity of the target molecule may be an activity associated with one or more biological effects of the target molecule on a biological system. For example, the desired activity of the target molecule may be associated with its ability to control the proliferation of neoplastic cells.

The desired activity may also be any activity that is associated with a particular region (or regions) of a target molecule. For example, the desired activity may be the activity associated with a region or regions of a drug molecule or a specific protein that has, or is likely to have, biological activity.

Preferably, the target molecule is a drug molecule. More preferably, the drug molecule is flurbiprofen or sulindac sulfide.

An example of an analogue of flurbiprofen that may be used as an active target molecule is the (*R*)-stereoisomer of a molecule with the following chemical formula:



(I)

or a salt thereof, wherein:

- R^1 is selected from hydrogen and lower alkyl (C1 to C8);
- R^2 is $YX^2((CH_2)_m X^2)_n$-, wherein m is 2 to 4, n is 1 to 6, X^2 is selected from O, S and N, and Y is independently selected from hydrogen, lower alkyl, or a suitable heteroatom protecting group;
- R^3 is selected from one or more of hydrogen, alkyl, aryl, halogen, hydroxy, alkoxy, aryloxy, amino (unsubstituted and substituted) and caroboxy;
- R^4 is selected from one or more of hydrogen, alkyl, aryl, halogen, hydroxy, alkoxy, aryloxy, amino (unsubstituted and substituted) and caroboxy;
- X is selected from fluoro, chloro, bromo and iodo;
- M is selected from hydroxy, alkoxy, aryloxy, amino, alkylamino (mono- and di-), arylamino (mono- and di-), N-morpholino, hydroxyalkylamino, dialkylaminoalkylamino, aminoalkylamino, polyhydroxyamino, and salts of any of the aforementioned.

X is preferably fluoro and most preferably substituted *meta* to the alkylcarboxylate group.

Preferably, R^1 is a lower alkyl group, and is most preferably a methyl group.

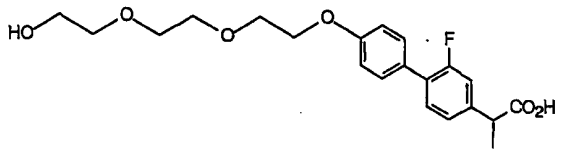
Preferably, R^2 is an alkyleneoxy or polyoxyalkylene chain, more preferably having between 1 and 4 alkyleneoxy repeating units. Suitable alkyleneoxy repeating units include ethyleneoxy and propyleneoxy. In one particularly preferred form of the invention, R^2 is a triethylene glycol group.

Preferably, the R^2O- group is substituted at a position *para* to the aryl substituent.

Preferably, both R^3 and R^4 are hydrogen.

M is preferably hydroxy or a salt thereof.

From the above, it will be evident that in one particularly preferred form, the present invention provides a compound of formula (III):

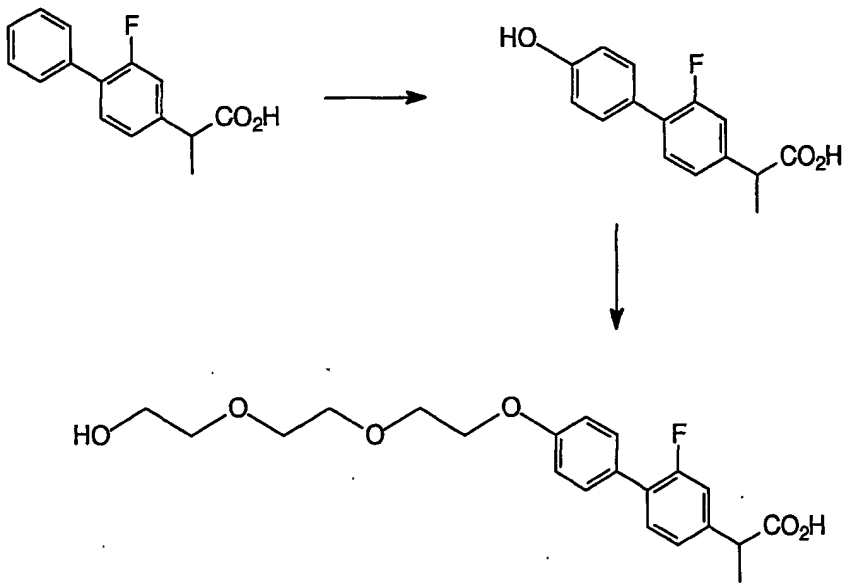


(III)

Compounds of formula (I) and (III) are available as (*R*)- and (*S*)-stereoisomers and the present invention contemplates the production and/or use of either pure (*R*)- or pure (*S*)-stereoisomers, as well as racemic mixtures or mixtures enriched with either stereoisomer. The (*R*)-stereoisomer of compounds of formula (I) or (III) may be active and therefore suitable as a target molecule in the methods of the present invention. Conversely, the (*S*)-stereoisomer of compounds of formula (I) or (III) may be inactive or have a reduced activity relative to the (*R*)-stereoisomer and therefore the (*S*)-stereoisomer may be suitable as a second molecule in the methods of the present invention.

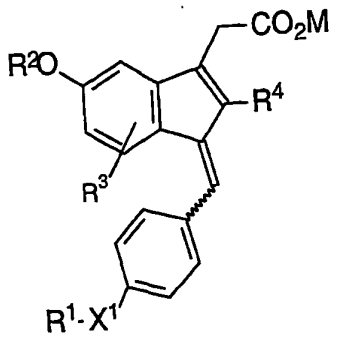
The stereoisomers of compounds of formula (I) may be separated by any of the techniques used for that purpose in the art, including chromatography for example.

Compounds of formula (I) may be produced by any suitable synthetic method, including those known methods for the production of flurbiprofen. An example of a suitable synthetic method is shown in Scheme 1, in which the method includes the steps of hydroxylation of the unsubstituted phenyl ring in flurbiprofen, followed by substitution of the free phenol with the triethylene glycol group.



Scheme 1

An example of an analogue of sulindac sulfide that may be used as an active target molecule is a molecule with the following chemical formula:



(II)

or a salt thereof, wherein:

- X^1 is selected from sulfide, sulfone and sulfoxide;
- R^1 is selected from hydrogen, hydroxy (when X^1 is sulfone or sulfoxide), and lower alkyl (C1 to C8);
- R^2 is $YX^2((CH_2)_m X^2)_n$-, wherein m is 2 to 4, n is 1 to 6, X^2 is selected from O, S and N, and Y is independently selected from hydrogen, lower alkyl, or a suitable heteroatom protecting group;
- R^3 is selected from hydrogen, halogen, alkyl, alkoxy, acyloxy, amino, alkylamino (mono- and di-), arylamino (mono- and di-), nitro, cyano, carboxyl;
- R^4 is selected from hydrogen and lower alkyl (C1 to C8); and
- M is selected from hydroxy, alkoxy, aryloxy, amino, alkylamino (mono- and di-), arylamino (mono- and di-), N-morpholino, hydroxyalkylamino, dialkylaminoalkylamino, aminoalkylamino, polyhydroxyamino, and salts of any of the aforementioned.

X^1 is preferably either a sulfone or a sulfide, and is most preferably a sulfide.

Preferably, R^1 is a lower alkyl group, and is most preferably a methyl group.

Preferably, R^2 is an alkyleneoxy or polyoxyalkylene chain, more preferably having between 1 and 4 alkyleneoxy repeating units. Suitable alkyleneoxy

repeating units include ethyleneoxy and propyleneoxy. In one particularly preferred form of the invention, R^2 is a triethylene glycol group.

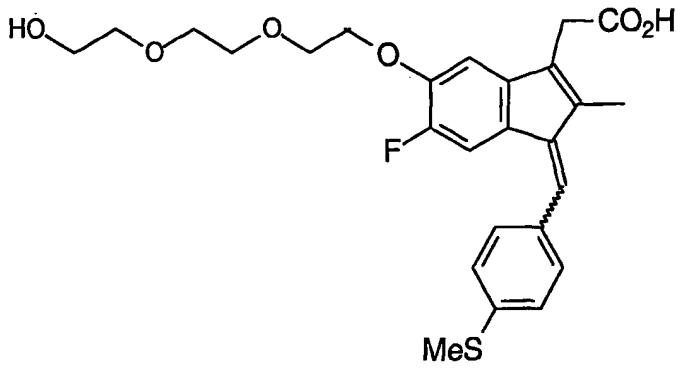
Preferably, R^3 is a halogen group (iodo-, bromo-, chloro- or fluoro-), more preferably a fluoro group, and most preferably a fluoro group ortho to the hydroxy group.

Preferably, R^4 is a lower alkyl group and is most preferably methyl.

M is preferably hydroxy or a salt thereof.

Compounds of formula (II) are available as (*E*) or (*Z*) geometric isomers. The compounds may be used as a mixture of (*E*) and (*Z*) isomers (whether a 1:1 mixture or some other ratio), or the geometric isomers may be separated using any of the standard techniques that are used for that purpose, such as chromatography.

From the foregoing it will be evident that in one preferred form, the invention provides a compound of formula (IV), or a salt thereof:



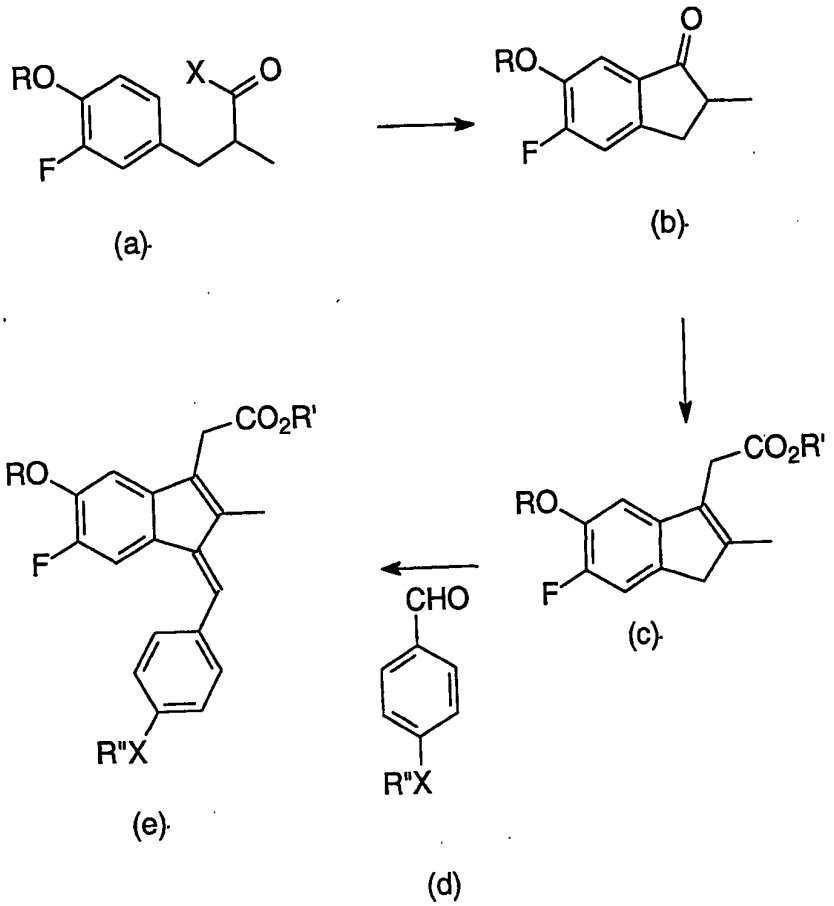
(IV)

It will be appreciated that compounds of formula (II) and (IV) may be formed by any one of a number of synthetic routes. However, in one form the present invention also provides a process for the preparation of compounds of formula (II), the process including the steps of:

i. intramolecular cyclisation of acyl donor (a) to form ketone (b);

ii. enolate addition to (b) followed by dehydration to form indene (c); and

iii. addition of (c) to aldehyde (d) to form indene (e).

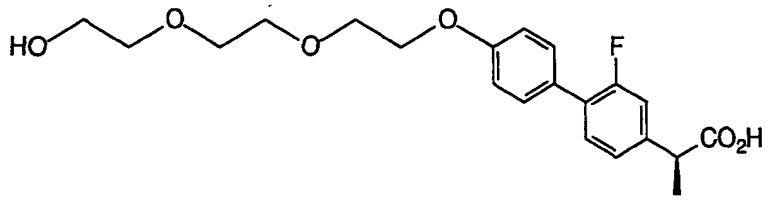


The second molecule according to the methods of the present invention is any molecule that is structurally similar to the target molecule, but which does not show a desired activity associated with the target molecule.

Structurally similar molecules include isomeric molecules such as isomers, geometric isomers, enantiomers, conformers, stereoisomers, structural isomers, molecules that substitute one or more chemical groups in a molecule with other chemical groups, or molecules that are substantially similar in the three dimensional structure of one or more parts of the molecule.

For example, the second molecule may include a drug analogue that shows reduced activity as compared to the biologically active drug. The second molecule may be an analogue of a known drug, such as flurbiprofen or sulindac sulfide. The second molecule may also be a molecule that is structurally similar to the target molecule, but which has been altered in a region (or regions) that is of interest in the target molecule.

An example of an analogue of flurbiprofen that may be utilised as a second molecule in the methods of the present invention is the molecule with the following chemical formula:



This molecule is the (*S*)-stereoisomer. The (*R*)-stereoisomer shows similar activity to flurbiprofen, while the (*S*)-stereoisomer is inactive.

The source of the pool of candidate proteins in the methods of the present invention may be any source of proteins, including proteins derived from cellular or viral extracts, proteins displayed on the surface of a viral particle by cloning one or more DNAs into a suitable viral vector, the protein products of one or more DNAs cloned into a protein expression vector and expressed in a suitable expression system, the products of *in vitro* translation of different mRNAs, or chemically synthesized polypeptides or proteins. The pool of candidate proteins may also contain post-translation modifications. For example, the pool of candidate proteins may include one or more glycosylated proteins.

Preferably, the source of the pool of candidate proteins is a pool of proteins expressed from suitable DNA molecules inserted into a viral genome. More preferably, the source of the pool of candidate proteins is a pool of viral particles

wherein each of the candidate proteins in the pool is displayed on the surface of a viral particle.

In the case of a pool of candidate proteins derived from cellular extracts, the source of the pools of candidate proteins includes cellular extracts derived from cell populations, group of cells, tissues or organs. Cellular extracts may be prepared by suitable methods known in the art. Cellular extracts may be derived from any prokaryotic or eukaryotic organism, including animals or humans.

In the case of cellular extracts derived from tissues, the cellular extract may be derived from cells selected from one or more of the following types of tissue: colorectal tissue, breast tissue, cervical tissue, uterine tissue, renal tissue, pancreatic tissue, esophageal tissue, stomach tissue, lung tissue, brain tissue, liver tissue, bladder tissue, bone tissue, prostate tissue, skin tissue, ovary tissue, testicular tissue, muscle tissue or vascular tissue.

These tissues may further contain cells that are normal (non-cancerous), pre-cancerous (having acquired some but not all of the cellular mutations required for a cancerous genotype) or cancerous cells (malignant or benign). Such tissues may contain cells that are normal, pre-cancerous or cancerous, any combination of cells that are normal, pre-cancerous or cancerous, or any other form of diseased cell. As will be readily appreciated, there are numerous methods well known in the art for determining whether cells are normal, pre-cancerous, cancerous or diseased, including histopathology and other phenotypic and genotypic methods of identifying cells.

As stated above, the source of the pool of proteins may also be a pool of proteins expressed from one or more DNAs inserted into a viral genome and which are displayed on the surface of the viral particle when expressed. The DNAs inserted may be complementary DNAs (cDNAs), DNA fragments derived from genomic or viral DNAs, or chemically synthesized DNAs. The DNAs inserted into the viral genome may also make up a library of DNAs, including libraries of cDNAs produced by reverse transcription of cellular mRNAs and

genomic DNA fragments. The DNAs inserted into the viral genome may also include one or more random DNA sequences, resulting in the expression of random polypeptides. For example, the random DNAs so inserted may be chemically synthesized DNAs.

Suitable viruses for cloning of DNA so as to express proteins on the surface of the viral particle include bacteriophage viruses such as those derived from T7, T4, lambda, lambdoid phage, or filamentous phage, including M13, f1 and fd. The DNAs may be cloned into the vectors derived from these phage by suitable methods known in the art. The insertion of the DNA to be expressed into an appropriate coat protein gene allows the protein encoded by the DNA to be displayed on the surface of a viral particle.

Viruses displaying proteins on their surface may then be produced by infection of a suitable host and preparation of viral extracts by suitable methods that are known in the art.

The source of the pool of proteins may also be a pool of proteins expressed from one or more DNAs inserted into a vector (plasmid vector or viral genome) and expressed in a suitable expression system. Once again, the DNAs inserted may be complementary DNAs (cDNAs), DNA fragments derived from genomic or viral DNAs, or chemically synthesized DNAs. The DNAs inserted into the vector may also make up a library of DNAs, including libraries of cDNAs and genomic DNA fragments. The DNAs inserted into the vector may also include one or more random DNA sequences, resulting in the expression of random polypeptides. For example, the random DNAs so inserted may be chemically synthesized DNAs.

The selectable moiety coupled to the non-nucleic acid target molecule according to the methods of the present invention is any moiety that allows the target molecule to be substantially purified away from other molecules, including the pool of candidate proteins. For example, the selectable moiety may be a chemical group such as an activated carbonate group that allows the target molecule to be covalently linked to a solid support. For protein molecules,

covalent coupling of the protein molecule to the solid support also includes coupling to the solid support via primary amines or cysteines by suitable methods that are known in the art.

In addition, a chemical moiety such as a biotin containing group may be coupled to the target molecule allowing the target molecule to be captured by an avidin or streptavidin group coupled to a solid support.

Further examples of methods of immobilisation of the target molecule include coupling of an antigen to the target molecule and capture by an immobilised antibody. Immobilisation of the target molecule may also utilise capture of glutathion-S-transferase-fusion proteins by anti GST antibodies, capture of 6xHis-fusion proteins by anti 6xHis antibodies or a nickel-chelating surface, capture of cAMP or cGMP binding proteins by cAMP or cGMP immobilised on a solid support. In these cases, the target molecule may be coupled to a protein molecule to be captured, or alternatively, the target molecule may be a protein engineered to be able to be captured by one of these methods.

The binding of the pool of candidate proteins to the non-nucleic acid target molecule in the presence of the second molecule in the methods of the present invention may be achieved under conditions suitable to the particular pool of candidate proteins being used. For example, the temperature and solution may be selected depending upon the properties of the pool of candidate proteins being used. Preferably, the temperature of binding may be within the range from 4°C to 42°C, and the binding achieved in a suitable buffer, including the use of tris-based and/or phosphate buffered solutions. Such solutions may further include appropriate amounts of further components, including salts, detergents and other agents depending upon the properties of the particular pool of candidate proteins being used.

The target molecule may be free in solution when binding occurs (and then subsequently captured), or alternatively be immobilised to a solid support during

the binding reaction. In a similar fashion, the second molecule may be free in solution, or alternatively, be immobilised to a solid support.

Preferably, the binding of the candidate proteins to the target molecule in the presence of the second molecule is performed under conditions where the second molecule is present in a molar excess to the target molecule. Most preferably the second molecule is in a molar excess of at least one hundred fold.

The ratio of the second molecule to the target molecule may also be selected so as to obtain binding proteins that bind to the target molecule with varying affinity. The greater the ratio of the second molecule to the target molecule in the binding reaction, the greater the affinity of proteins that bind to the target is likely to be.

In the methods of the present invention, the proteins bound to the target molecule may be isolated by a suitable means. If the target molecule is coupled to a fixed solid support, the proteins bound to the target molecule may be isolated by washing the solid support in a suitable buffer to remove any proteins that do not bind to the target molecule. If the target molecule is coupled to a solid support such as beads (for example paramagnetic beads), the beads may first be isolated and the proteins that do not bind removed by washing the beads in a suitable buffer. Alternatively, the target molecule may be captured and thus immobilised on the solid support, and proteins that do not bind them removed by washing.

The removal of proteins bound to the target molecule may also be achieved by eluting in a suitable buffer containing free target molecule in substantial molar excess. Preferably, the free target molecule is in a molar excess of greater than one thousand fold.

In removing proteins that are bound to the target molecule, the eluates may be isolated at different times during the washing procedure to allow for the identification of binding proteins that have different dissociation rates.

The proteins so isolated in the methods of the present invention may be subject to methods to allow their identification and/or characterisation. For example, if the proteins isolated are sufficiently pure, the amino acid sequence of a binding protein isolated according to the present invention may be determined by a method that includes determination of the amino acid sequence. Alternatively, mass spectrometry may be performed.

In the case of viral particles that express a candidate protein on their surface, the identity of the expressed protein may be determined directly by determination of the DNA sequence of the DNA inserted into the viral genome that results in expression of the protein on the surface of the virus. As will be appreciated, determination of the DNA sequence will allow the prediction of the amino acid sequence of the protein that is displayed on the surface of the viral particle.

In addition, the proteins so isolated that bind to the target molecule may be re-bound to the target molecule in the presence of the second molecule. In this way the process may be re-iterated, until a desired population of proteins that bind to the target molecule is achieved.

Confirmation of the ability of a binding protein isolated by the methods of the present invention to bind to the target molecule may be by a suitable method known in the art. For example, the protein may be substantially purified and then the binding affinity of the isolated protein to the target molecule be determined by a suitable means, including binding assays, surface plasmon resonance or atomic force microscopy. Alternatively, the binding affinity of a phage particle displaying the binding protein to the drug may be determined by a suitable means, including surface plasmon resonance or atomic force microscopy.

In another form, the present invention provides a method for identifying a protein capable of binding to a target molecule, the method including the steps of:

(a) providing a pool of candidate proteins, wherein each candidate protein is displayed on the surface of a viral particle;

(b) providing a non-nucleic acid target molecule, wherein the non-nucleic

acid target molecule is coupled to a selectable moiety;

(c) providing a second molecule which is structurally similar to the non nucleic acid target molecule, wherein the second molecule is deficient in a desired activity of the target molecule;

(d) allowing one or more of the candidate proteins to bind to the non nucleic acid target molecule in the presence of the second molecule;

(e) isolating one or more proteins bound to the target molecule;

(f) amplifying the viral particles encoding the isolated binding proteins;

(g) reiterating steps (a) through (f); and

(h) identifying the binding protein.

In this form of the present invention, the source of the pool of proteins is a pool of proteins expressed from one or more DNAs inserted into a viral genome, the proteins so expressed being displayed on the surface of the viral particle. Suitable viruses for cloning of DNA so as to express proteins on the surface of the viral particle include bacteriophage viruses such as those derived from T7, T4, lambda, lambdoid phage, or filamentous phage, including M13, f1 and fd. The DNAs may be cloned into vectors derived from these phage by suitable methods known in the art.

The DNAs inserted into such viral vectors may be complementary DNAs (cDNAs), DNA fragments derived from genomic or viral DNAs, or chemically synthesized DNAs. The DNAs inserted into the viral genome may also make up a library of DNAs, including libraries of cDNAs (for example obtained by the reverse transcription of cellular mRNAs) and genomic DNA fragments. The DNAs inserted into the viral genome may also include one or more random DNA sequences, resulting in the expression of random polypeptides. For example, the random DNAs so inserted may be chemically synthesized DNAs.

Viruses displaying proteins on their surface may then be produced by infection of a suitable host and preparation of viral extracts by methods that are well known in the art. In this way, a pool of candidate proteins in which each candidate protein is displayed on the surface of the viral particle may be produced.

The amplification of viral particle encoding the isolated proteins may be achieved by re-infecting a competent viral host with the isolated viral particles by a suitable procedure. The viral particles so concentrated may be concentrated by a suitable means, so as to allow the process of binding and isolating the proteins that bind to the target molecule to be reiterated.

Preferably, a proportion of the DNA inserts that make up the viral population after amplification may be characterised for insert size and their DNA sequence. For example, the DNA inserted into each viral particle may be isolated by obtaining a pure viral population by way of an isolated plaque and isolating the DNA inserted in that particular viral DNA by polymerase chain reaction using appropriate primers. The size of the DNA inserts may be determined by a suitable method. The DNA sequence of the DNA insert may be determined by a suitable method.

The reiteration step of this form of the present invention is continued until a desired level of representation of the DNA inserts is reached in the viral population. The representation of the DNA inserts in the viral population may be determined by a suitable method. Preferably five or more reiterations are performed.

The identity of the binding protein may then be determined by determination of the DNA sequence of the DNA inserted into the viral genome that results in expression of the protein on the surface of the virus. As will be appreciated, determination of the DNA sequence will allow the prediction of the amino acid sequence of the protein expressed on the surface of the viral particle.

The present invention also provides a method for identifying a protein capable of binding to target molecule, the method including the steps of:

(a) providing a first pool of candidate proteins;

(b) providing a non-nucleic acid target molecule, wherein the non-nucleic acid target molecule is coupled to a selectable moiety;

(c) providing a second molecule which is structurally similar to the non-nucleic acid target molecule, wherein the second molecule is deficient in a desired activity of the target molecule;

(d) allowing one or more of the candidate proteins in the first pool to bind to the non-nucleic acid target molecule in the presence of the second molecule;

(e) isolating a protein in the first pool that binds to the target molecule;

(f) comparing the level of the protein in the first pool of candidate proteins with the level of the protein in a second pool of proteins; and

(g) identifying a protein that is differentially represented between the first and second pools.

In this form of the present invention, the identification of proteins that bind to a target molecule and which are differentially represented between two pools of candidate proteins may be achieved.

The determination of the level of a protein in the first pool may be achieved by a suitable procedure known in the art, including the determination of the concentration by methods that include the use of antibodies to detect the binding protein. For example, the concentration of the binding protein in the first pool may be achieved with an antibody raised to the binding protein, and the subsequent use of the antibody to visualise the protein by Western analysis or the use of the antibody to immunoprecipitate the protein.

The level of the binding protein in a second pool of candidate proteins may then be determined in a similar fashion. In this way, proteins isolated from the first pool of candidate proteins may be compared with a second pool of candidate proteins, so as to identify proteins that are differentially represented between the two pools of binding proteins.

For example, the first pool of proteins may be derived from a tissue that contain cells that are normal (non-cancerous), and the second pool of proteins may be derived from cells that are pre-cancerous (having acquired some but not all of

the cellular mutations required for a cancerous genotype) or cancerous cells (malignant or benign). Any differences in the level of a binding protein between the pool of proteins derived from a normal tissue and another tissue allows the identification of binding proteins that are differentially expressed between the two different pools of proteins.

The present invention also provides a method for identifying a protein capable of binding to target molecule, the method including the steps of:

(a) providing first and second pools of candidate proteins;

(b) providing a non-nucleic acid target molecule, wherein the non-nucleic acid target molecule is coupled to a selectable moiety;

(c) providing a second molecule which is structurally similar to the non-nucleic acid target molecule, wherein the second molecule is deficient in a desired activity of the target molecule;

(d) allowing one or more of the candidate proteins in the first pool to bind to the non-nucleic acid target molecule in the presence of the second molecule;

(e) isolating one or more proteins in the first pool that bind to the target molecule;

(f) allowing one or more of the candidate proteins in the second pool to bind to the non-nucleic acid target molecule in the presence of the second molecule;

(g) isolating one or more proteins in the second pool that bind to the target molecule; and

(h) comparing the level of one or more proteins isolated from each of the first and second pools to identify a protein that is differentially represented between the first and second pools.

In this form of the present invention, the identification of proteins that bind to a target molecule and which are differentially represented between two pools of candidate proteins may also be achieved.

Preferably, the reactions utilising the first and second pools are performed in parallel experiments under exactly the same conditions.

The determination of the level of a protein isolated from the first pool may be achieved by a suitable procedure, including the determination of the concentration by methods that include the use of antibodies to detect the binding protein. For example, the concentration of the binding protein isolated from the first pool of candidate proteins may be achieved with an antibody raised to the binding protein, and the subsequent use of the antibody to visualise the protein by Western analysis or the use of the antibody to immunoprecipitate the protein.

The level of the binding protein isolated from a second pool of candidate proteins may then be determined in a similar fashion. In this way, proteins isolated from the first pool of candidate proteins may be compared with those isolated from a second pool of candidate proteins, so as to identify isolated proteins that are differentially represented between the two pools of binding proteins.

For example, the first pool of proteins may be derived from a tissue that contain cells that are normal (non-cancerous), and the second pool of proteins may be derived from cells that ate pre-cancerous (having acquired some but not all of the cellular mutations required for a cancerous genotype) or cancerous cells (malignant or benign). Any differences in the level of an isolated binding protein between the protein isolated from a normal tissue and that isolated from another tissue allows the identification of binding proteins that are differentially represented between the two different pools of proteins.

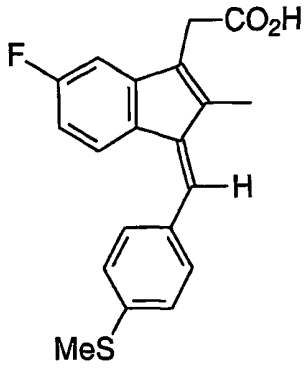
Description of the Preferred Embodiments

The present invention will now be described in relation to various examples of preferred embodiments. However, it must be appreciated that the following description is not to limit the generality of the above description.

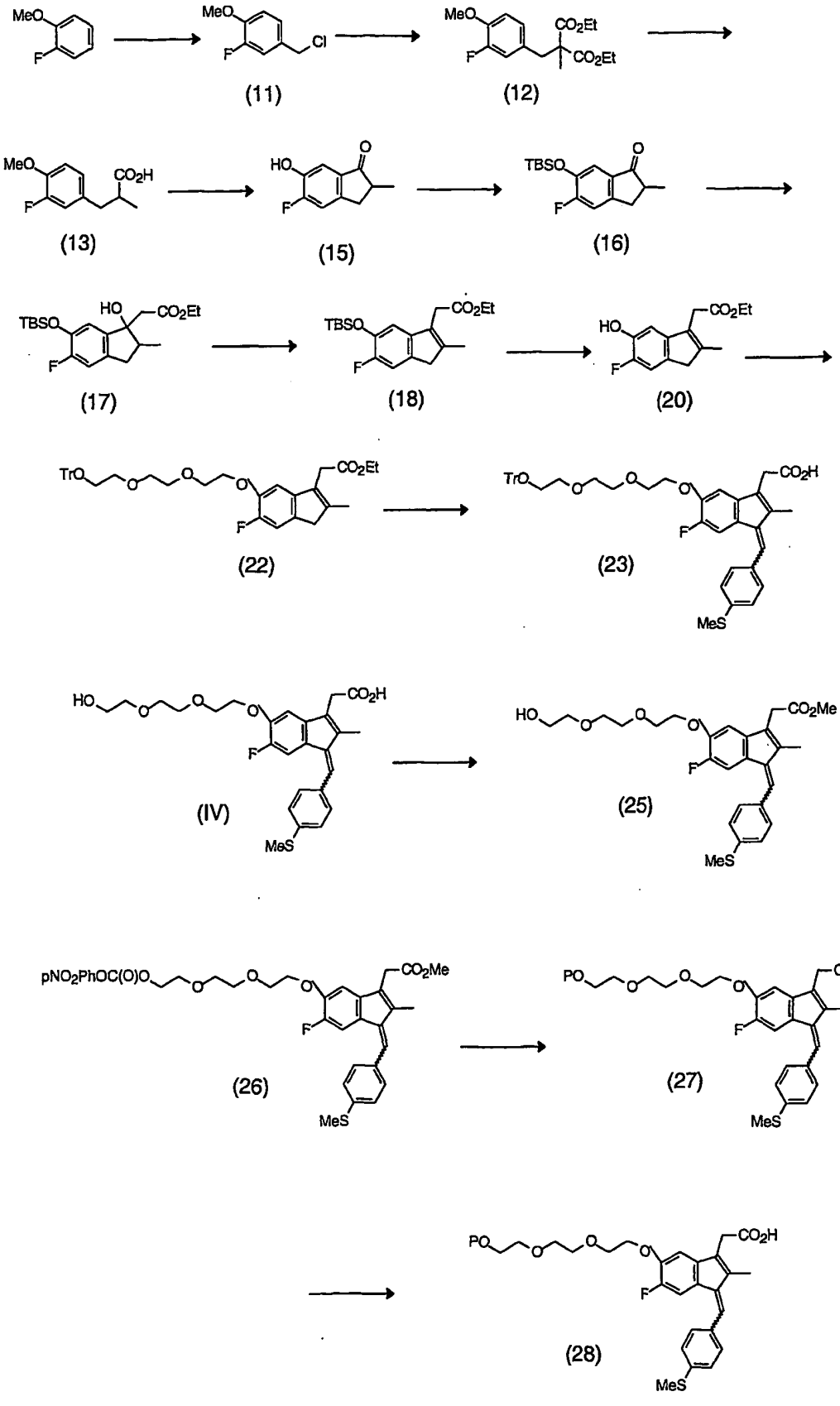
Example 1 - Synthesis of analogues of sulindac sulfide and attachment to a solid phase

Sulindac sulfide is a drug that acts to decrease the number of precancerous lesions (adenomas) in the colon both in animals and in humans.

The structure of sulindac sulfide is as follows:



An analogue of sulindac sulfide, designated as compound (IV), was prepared from 2-fluoroanisole as shown in Scheme 2.


MeO
F
Cl
CO2Et
CO2H
HO
O
TBSO
HO
TrO
pNO2PhOC(O)O
PO
CO2Me
CO2
MeS
(11)
(12)
(13)
(15)
(16)
(17)
(18)
(20)
(22)
(23)
(IV)
(25)
(26)
(27)
(28)

Scheme 2

Example 1.1 – Synthesis of 4-(chloromethyl)-2-fluoro-1-methoxybenzene (11)

Anhydrous hydrogen chloride (excess) was bubbled into a stirred mixture of 2-fluoroanisole (16.78 g, 0.133 mol), paraformaldehyde (8.39 g, 0.279 mol), anhydrous powdered zinc chloride (1.45 g) and acetic acid (65 cm^3) at 40°. After *ca.* 1 h the hydrogen chloride stream was stopped and the reaction stirred at 40° for a further 8 h. During this time the mixture became homogeneous. Upon cooling, water (50 cm^3) and toluene (50 cm^3) were added and the resultant layers separated. The water layer was extracted with more toluene (20 cm^3), the combined toluene extracts were washed with aqueous 2M sodium hydroxide (2x75 cm^3) and finally with saturated aqueous sodium chloride. The toluene solution was dried ($MgSO_4$), filtered and evaporated *in vacuo* to give the chloride (11) as a colourless oil (22.68 g, 97%). ^{1}H-NMR ($CDCl_3$, 300 MHz): δ 3.89 (s, 3H); 4.53 (s, 2H); 6.92 (m, 1H); 7.11 (m, 2H).

Example 1.2 – Synthesis of diethyl 2-[(3-fluoro-4-methoxyphenyl)methyl]-2-methylpropane-1,3-dioate (12)

Diethyl methylmalonate (27.87 g, 0.160 mol) was added to a stirred solution of sodium ethoxide (0.160 mol) in anhydrous ethanol (140 cm^3) at room temperature. 3-Fluoro-4-methoxybenzyl chloride (11) was then added to this solution over a 0.5 h period. The mixture was then refluxed for 16 h, followed by cooling and removal of most of the ethanol. The residue was taken up into water (100 cm^3) and ether (100 cm^3). The phases were separated and the aqueous layer washed with more ether. The combined ether washings were shaken with saturated aqueous sodium chloride. The ethereal solution was dried ($MgSO_4$), filtered and evaporated *in vacuo* to give the diester (12) as a colourless oil (45.29 g, 91%). ^{1}H-NMR ($CDCl_3$, 300 MHz): δ 1.25 (t, *J* 7 Hz, 6H); 1.33 (s, 3H); 3.14 (s, 2H); 3.85 (s, 3H); 4.19 (q, *J* 7 Hz, 4H); 6.85 (m, 3H).

Example 1.3 – Synthesis of 3-(3-fluoro-4-methoxyphenyl)-2-methylpropanoic acid (13)

The diester (12) was added to a solution of potassium hydroxide (40.00 g, 0.713 mol) in water (60 cm^3) and ethanol (50 cm^3) and the mixture refluxed for 16 h. Upon cooling, a mixture of concentrated sulfuric acid (70 g) and water (50 cm^3) was added carefully and reflux recommenced. After 15 h, the reaction mixture was cooled and extracted with ether three times. The combined ether washings were shaken with saturated aqueous sodium chloride then dried ($MgSO_4$), filtered and evaporated *in vacuo* to give the acid (13) as a pale yellow oil (27.13 g, 89%). ^{1}H-NMR ($CDCl_3$, 300 MHz): δ 1.78 (d, *J* 7 Hz, 3H); 2.68 (m, 2H); 2.98 (dd, *J* 6 Hz, 13 Hz, 1H); 3.87 (s, 3H); 6.90 (m, 3H).

Example 1.4 – Synthesis of 5-fluoro-6-methoxy-2-methylindan-1-one (14)

The acid (13) (26.95 g, 0.127 mol) was added to polyphosphoric acid (270 g) at room temperature. The mixture was swirled by hand in an oil bath heated to 90° until homogeneity was achieved. The viscous solution was stirred magnetically for a further 2 h at 90°. The hot mixture was then poured on to crushed ice (500 g), ether (100 cm^3) was added and the mixture stirred at room temperature for 15 h. The layers were separated and the aqueous layer washed with more ether (2x50 cm^3). These ethereal washings were combined with the original ether layer and washed with saturated aqueous sodium bicarbonate (50 cm^3) followed by saturated aqueous sodium chloride. Drying ($MgSO_4$), filtration and removal of solvent gave the indanone as a white solid (23.32 g, 95%). ^{1}H-NMR ($CDCl_3$, 300 MHz): δ 1.30 (d, *J* 7.5 Hz, 3H); 2.69 (m, 2H); 3.31 (dd, *J* 7.5 Hz, 17 Hz, 1H); 3.91 (s, 3H); 7.12 (dt, *J* 10.5 Hz, 1 Hz, 1H); 7.29 (d, *J* 8 Hz, 1H).

Example 1.5 – Synthesis of 5-fluoro-6-hydroxy-2-methylindan-1-one (15)

The methoxyindanone (14) (23.00g, 0.118 mol) and tetrabutylammonium bromide (3.80 g, 11.8 mmol) were dissolved in 48% aqueous hydrogen bromide

(130 ml) and the stirred solution was heated at 115° for 5.5 h. Upon cooling, water (300 cm^3) and ether (200 cm^3) were added and the resultant layers separated. The aqueous layer was washed with more ether (2x50 cm^3) and these washings were combined with the original ether layer and extracted with 5% aqueous sodium hydroxide (2x100 cm^3). The ether layer was discarded and the alkaline aqueous layer acidified to pH 1 with 50% aqueous sulfuric acid. Extraction with ether, drying ($MgSO_4$), filtration and evaporation of the ether *in vacuo* afforded the crude hydroxyindanone (20.55g) as a dark brown solid. "Dry Column" Flash Chromatography (13 cm diameter sintered glass funnel, 7 cm depth of flash silica, hexane/ethyl acetate gradient) gave the pure hydroxyindanone (15) (15.28 g, 72%) as a pale yellow solid. ^{1}H-NMR ($CDCl_3$, 300 MHz): δ 1.30 (d, *J* 7.5 Hz, 3H); 2.70 (m, 2H); 3.31 (dd, *J* 7.5 Hz, 16.5 Hz, 1H); 5.98 (broad s, 1H); 7.14 (d, *J* 10 Hz, 1H); 7.38 (d, *J* 8 Hz, 1H). Later fractions yielded the hydroxyindanone in lower purity: 2.37g (11%) of yellow solid pure enough (^{1}H-NMR) for the next step and finally 2.33 g (11%) of a dark brown oil insufficiently pure (^{1}H-NMR) for further use.

Example 1.6 – Synthesis of 5-fluoro-2-methyl-6-(tert-butyldimethylsilyloxy)indan-1-one (16)

The hydroxyindanone (15) (9.01 g, 50.0 mmol) and imidazole (8.51 g, 0.125 mol) were dissolved in dry dimethylformamide (40 cm^3). *Tert*-Butyldimethylsilyl chloride (9.04 g, 60 mmol) was added to this solution and stirring was continued at room temperature for 16 h. The reaction mixture was combined with 5% aqueous sodium bicarbonate (210 cm^3) and extracted with hexane (75 cm^3 then 2x30 cm^3). The combined hexane extracts were dried ($MgSO_4$), filtered and the solvent evaporated *in vacuo* to yield a pale yellow oil (15.30 g) which solidified upon standing at room temperature. ^{1}H-NMR analysis indicated that the product was a mixture containing the desired silyl ether (16) (*ca.* 86 mole %) and two unidentified components with just *tert*-butyldimethylsilyl ^{1}H-NMR signals (total *ca.* 14 mole %). ^{1}H-NMR of the silyl ether ($CDCl_3$, 300 MHz): δ 0.20 (s, 6H);

1.00 (s, 9H); 1.29 (d, *J* 7 Hz, 3H); 2.68 (m, 2H); 3.31 (dd, *J* 7.5 Hz, 16.5 Hz, 1H); 7.10 (d, *J* 10 Hz, 1H); 7.25 (d, *J* 8 Hz, 1H). Without purification, this crude product was successfully used in the next step.

Example 1.7 – Synthesis of ethyl 2-[5-fluoro-1-hydroxy-2-methyl-6-(tert-butyldimethylsilyloxy)indanyl]acetate (17)

A 1.0 M solution of lithium hexamethyldisilazide in tetrahydrofuran (50 cm^3, 50 mmol) was cooled to -75° by stirring in a dry ice/acetone bath. Ethyl acetate (4.9 cm^3, 50 mmol) was added over a period of 3 minutes and stirred at -75° for a further 15 minutes. A solution of the crude ketone (16) (15.30 g, *ca.* 50 mmol) in dry tetrahydrofuran (30 cm^3) was added to the lithium enolate at a rate slow enough to keep the internal reaction temperature below -60° (*ca.* 20 minutes). After stirring at -75° for a further 5 minutes, 20% hydrochloric acid (10 cm^3) was added. After the mixture had warmed to room temperature, the tetrahydrofuran was removed and ether (100 cm^3) and water (50 cm^3) were added. The layers were separated and the aqueous phase washed with more ether (50 cm^3). The combined ether solutions were washed with saturated aqueous sodium chloride. Drying ($MgSO_4$), filtration and removal of solvent gave the alcohol (17) (18.52 g, 97%) as a yellow-orange oil. ^{1}H-NMR ($CDCl_3$, 300 MHz): δ 0.17 (s, 6H); 0.99 (s, 9H); 1.08 (d, *J* 6.5 Hz, 3H); 1.28 (t, *J* 7 Hz, 3H); 2.38 (m, 1H); 2.56 (m, 1H); 2.72 (m, 2H); 2.91 (dd, *J* 7.5 Hz, 15.5 Hz, 1H); 4.20 (q, *J* 7 Hz, 2H); 6.87 (m, 2H).

Example 1.8 – Synthesis of ethyl 2-[6-fluoro-2-methyl-5-(tert-butyldimethylsilyloxy)inden-3-yl]acetate (18)

Sicapent™ (Merck, 17.8 g; 80% diphosphorus pentoxide, 13.4 g, 94.2 mmol) was added to a solution of the alcohol (17) (18.00 g, 47.1 mmol) in benzene. The mixture was refluxed for 0.5 h, cooled and filtered through flash silica. Elution with more benzene (50 cm^3) was followed by elution with ether until the eluate was colourless. The benzene and ether solutions were pooled and evaporated *in vacuo* to give a mixture of dehydration products (16.29 g, 95%) as a yellow-

orange oil. ^{1}H-NMR indicated that the indene (18) was the major product (*ca.* 75%). ($CDCl_3$, 300 MHz): δ 0.18 (s, 6H); 1.01 (s, 9H); 1.24 (t, *J* 7 Hz, 3H); 2.09 (s, 3H); 3.25 (s, 2H); 3.45 (s, 2H); 4.13 (q, *J* 7 Hz, 2H); 6.80 (d, *J* 8 Hz, 1H); 7.05 (d, *J* 11 Hz, 1H). The remaining 25% of the mixture consisted of the (*E*) and (*Z*) isomers of the corresponding *exo* alkene. Acid catalysed desilylation in the next step also isomerised this material to the desired indene.

Example 1.9 – Synthesis of 2-(6-fluoro-5-hydroxy-2-methylinden-3-yl)acetic acid (19)

The silyl ether (18) was added to ethanol (200 cm^3) which had been pre-treated with acetyl chloride (4.3 g, 55 mmol) and the solution refluxed for 3 hours. Upon cooling, the volatile components of the mixture were removed *in vacuo* and the residue was taken up into ether (100 cm^3). The ether solution was extracted with 1 M aqueous sodium hydroxide solution (50 cm^3 then 25 cm^3) and then discarded. Acidification of the aqueous washings with 25% w/w aqueous sulfuric acid was followed by extraction with ether (1x70 cm^3 then 2x30 cm^3). The combined ether washings were dried ($MgSO_4$), filtered and the ether evaporated *in vacuo* to give the phenolic carboxylic acid (19) (9.12 g, 94%) as a beige solid. ^{1}H-NMR ($CDCl_3$, 300 MHz): δ 2.09 (s, 3H); 3.26 (s, 2H); 3.50 (s, 2H); 6.88 (d, *J* 8 Hz, 1H); 7.08 (d, *J* 10 Hz, 1H).

Example 1.10 – Synthesis of ethyl 2-(6-fluoro-5-hydroxy-2-methylinden-3-yl)acetate (20)

The acid (19) (9.12 g, 41.0 mmol) was added to a solution of concentrated sulfuric acid (2.00 g, 20.4 mmol) in ethanol (100 cm^3). The solution was refluxed for 3.5 h, cooled and the ethanol removed *in vacuo*. The residue was taken up into ether (100 cm^3) and the solution washed with water (3x50 cm^3) followed by saturated aqueous sodium chloride. The ether solution was dried ($MgSO_4$), filtered and the ether evaporated *in vacuo* to give the ester (20) (9.75 g, 95%) as

a beige solid. ^{1}H-NMR ($CDCl_3$, 300 MHz): δ 1.25 (t, *J* 7 Hz, 3H); 2.09 (s, 3H); 3.25 (s, 2H); 3.46 (s, 2H); 4.14 (q, *J* 7 Hz, 2H); 6.89 (d, *J* 8 Hz, 1H); 7.07 (d, *J* 10 Hz, 1H).

Example 1.11 – Synthesis of 2-{2-[2-(triphenylmethoxy)ethoxy]ethoxy}ethan-1-ol (21)

A solution of trityl chloride (13.94 g, 50.0 mmol) in dichloromethane (25 cm^3) was added to a stirred solution of tri(ethylene glycol) (15.02 g, 0.100 mol) and triethylamine (7.69 g, 75.0 mmol) in dichloromethane (50 cm^3) at 0°. Stirring at 0° was continued for 2 h and then at room temperature for a further 14 h. The mixture was washed with 1M hydrochloric acid (50 cm^3), followed by water (50 cm^3) and saturated aqueous sodium chloride (50 cm^3). Drying ($MgSO_4$), filtration and evaporation of dichloromethane *in vacuo* gave an orange-brown oil which was purified by "Dry Column" Flash Chromatography (13 cm diameter sintered glass funnel, 7 cm depth of flash silica, hexane/ethyl acetate gradient) gave pure mono-tritylated material (21) (10.91 g, 56%) as a colourless oil. ^{1}H-NMR ($CDCl_3$, 300 MHz): δ 3.30 (t, *J* 5 Hz, 2H); 3.71 (m, 10H); 7.31 (m, 9H); 7.52 (m, 6H).

Example 1.12 – Synthesis of ethyl 2-[6-fluoro-2-methyl-5-(2-{2-[2-(triphenylmethoxy)ethoxy]ethoxy}ethoxy)inden-3-yl]acetate (22)

Diethyl azadicarboxylate (3.05 g, 17.5 mmol) was added slowly to a stirred solution of the alcohol (21) (6.26 g, 15.9 mmol), the phenol (20) (3.99 g, 15.9 mmol) and triphenylphosphine (4.60 g, 17.5 mmol) in dry tetrahydrofuran (70 cm^3) at 0°. Stirring was continued at 0° for 1 h and then for a further 66 h at room temperature, at which time TLC analysis indicated all of the phenolic starting material had been consumed. The tetrahydrofuran was removed *in vacuo* and the residue taken up into dichloromethane (50 cm^3) and loaded on to a bed of flash silica (6 cm depth, 10 cm diameter glass sinter funnel). "Dry Column" Flash Chromatography (hexane/ethyl acetate gradient) gave the pure

ether (22) (9.35 g, 94%) as a colourless oil. ^{1}H-NMR ($CDCl_3$, 300 MHz): δ 1.26 (t, *J* 7 Hz, 3H); 2.14 (s, 3H); 3.29 (t, *J* 5 Hz, 2H); 3.29 (s, 2H); 3.50 (s, 2H); 3.78 (m, 6H); 3.95 (t, *J* 5 Hz, 2H); 4.16 (q, *J* 7 Hz, 2H); 4.26 (t, *J* 5 Hz, 2H); 6.96 (d, *J* 7.5 Hz, 1H); 7.12 (d, *J* 11 Hz, 1H); 7.30 (m, 9H); 7.51 (m, 6H).

Example 1.13 – Synthesis of 2-[6-fluoro-2-methyl-1-[(4-methylthiophenyl)methylene]-5-(2-{2-[2-(triphenylmethoxy)ethoxy]ethoxy}ethoxy)inden-3-yl]acetic acid (23)

A *ca.* 0.5M solution of sodium methoxide in methanol was made by adding sodium (229 mg, 9.52 mmol) to dry methanol (20 cm^3). The indene (22) (2.98 g, 4.76 mmol) and 4-methylthiobenzaldehyde (797 mg, 5.24 mmol) were dissolved in this solution by swirling of the flask by hand. The resultant bright purple solution was refluxed for 1 h. During this time the solution turned orange and an orange oil separated out. Water (20 cm^3) was added and reflux continued for a further 0.5 h and the mixture became homogeneous. Water (250 cm^3) and ether (100 cm^3) were added to the cooled orange solution and the resultant emulsion broken by the addition of sodium chloride. The ether layer was discarded and the orange oil which had precipitated from the aqueous phase was dissolved with hot water. The aqueous solution was acidified with acetic acid (4 cm^3, 70 mmol) and extracted with ether (1x100 cm^3 then 2x50 cm^3). The combined ether washings were dried (Na_2SO_4), filtered and evaporated to dryness to give 2.60 g (75%) of the product (23) as a viscous orange oil. ^{1}H-NMR ($CDCl_3$, 300 MHz) revealed the presence of two geometric isomers in a *ca.* 5:1 ratio. Data for the major isomer (*Z*): δ 2.21 (s, 3H); 2.59 (s, 3H); 3.29 (t, *J* 5 Hz, 2H); 3.50 (s, 2H); 3.76 (m, 6H); 3.92 (t, *J* 5 Hz, 2H); 4.29 (t, *J* 5 Hz, 2H); 6.94 (d, *J* 8 Hz, 1H); 7.14 (s, 1H); 7.30 (m, 12H); 7.51 (m, 8H).

Example 1.14 – Synthesis of 2-(6-fluoro-5-{2-[2-(2-hydroxyethoxy)ethoxy]ethoxy}-2-methyl-1-[(4-methylthiophenyl)methylene]inden-3-yl)acetic acid (IV)

The trityl ether (23) (2.60 g, 3.56 mmol) was dissolved in a mixture of formic acid (40 cm^3) and ether (40 cm^3). After 16 h at room temperature, the solvents were evaporated *in vacuo* and the residue was taken up into ether (50 cm^3) and 0.2 M aqueous sodium hydroxide (50 cm^3). The layers were separated and the aqueous phase extracted with more ether (2x30cm^3), then acidified with 1 M hydrochloric acid. The acidified mixture was extracted with ether (3x50cm^3) and the combined ether extracts were dried (Na_2SO_4), filtered and the ether evaporated *in vacuo* to give the crude product (1.56 g) as an orange solid. Recrystallisation from ethyl acetate/hexane gave the pure (*Z*) isomer (IV) (0.823 g, 45%) as orange needles. ^{1}H-NMR ($CDCl_3$, 600 MHz): δ 2.18 (s, 3H); 2.55 (s, 3H); 3.58 (s, 2H); 3.63 (t, *J* 5 Hz, 2H); 3.68 (m, 4H); 3.76 (t, *J* 5 Hz, 2H); 3.81 (t, *J* 5 Hz, 2H); 4.30 (t, *J* 5 Hz, 2H); 7.00 (d, *J* 8 Hz, 1H); 7.10 (s, 1H); 7.18 (d, *J* 12.5 Hz, 1H); 7.30 (d, *J* 8 Hz, 2H); 7.43 (d, *J* 8 Hz, 2H).

This compound was used as an example of the chemistry that may be used to attach a drug to a solid phase.

To attach analogue (IV) to a solid phase, the alcohol group was converted to the corresponding 4-nitrophenyl carbonate. This active carbonate was reacted with TentaGel S-NH_2 to yield the polymer-supported derivative (27).

Example 1.15 – Synthesis of methyl 2-(6-fluoro-5-{2-[2-(2-hydroxyethoxy)ethoxy]ethoxy}-2-methyl-1-[(4-methylthiophenyl)methylene]inden-3-yl)acetate (25)

Material isolated from the mother liquor of recrystallisation of the acid (IV) (1.23 g, 2.52 mmol) was dissolved in methanol (40 cm^3). Sulfuric acid (200 mg, 2.04 mmol) was added and the solution was refluxed for 3 h. Upon cooling, the methanol was removed, ether (30 cm^3) and water (30 cm^3) were added and the phases separated. The ether phase was washed with saturated aqueous sodium bicarbonate solution (2x20 cm^3) then saturated aqueous sodium chloride solution (20 cm^3). Drying (Na_2SO_4), filtration and removal of the solvent *in*

vacuo gave 1.22 g of crude product. This material was purified by radial chromatography ("Chromatotron") using an ethyl acetate/hexane gradient to give 539 mg of an orange solid which still contained some of the minor (*E*) isomer. Recrystallisation (ethyl acetate/hexane) gave the pure (*Z*) isomer (25) (296 mg, 23%) as orange plates. ^{1}H-NMR ($CDCl_3$, 300 MHz): δ 2.17 (s, 3H); 2.55 (s, 3H); 3.56 (s, 2H); 3.66 (m, 2H); 3.72-3.82 (m, 6H); 3.74 (s, 3H); 3.88 (t, *J* 5 Hz, 2H); 4.23 (t, *J* 5 Hz, 2H); 6.84 (d, *J* 8 Hz, 1H); 7.10 (s, 1H); 7.24 (d, *J* 12 Hz, 1H); 7.29 (d, *J* 8 Hz, 2H); 7.43 (d, *J* 8 Hz, 2H).

Example 1.16 – Synthesis of methyl 2-[6-fluoro-2-methyl-1-[(4-methylthiophenyl)methylene]-5-(2-{2-[2-(4-nitrophenoxycarbonyloxy)ethoxy]ethoxy}ethoxy)inden-3-yl]acetate (26)

4-Nitrophenyl chloroformate (45 mg, 0.220 mmol) was added in one portion to a stirred solution of the alcohol (25) (98 mg, 0.195 mmol), *N*-methylmorpholine (30 mg, 30 mmol) and dimethylaminopyridine (1.2 mg, 10 mmol) in dichloromethane (5 cm^3) at 0°. The ice bath was removed and the mixture was stirred for 15 h. The dichloromethane was removed *in vacuo* and ethyl acetate and 1 M hydrochloric acid were added. The layers were separated and the organic phase washed with more 1 M hydrochloric acid, followed by saturated aqueous sodium hydrogen carbonate and finally with saturated aqueous sodium chloride. Drying ($MgSO_4$), filtration and removal of solvent *in vacuo* gave the crude product which was purified by Flash Chromatography (3:2 ethyl acetate/hexane) to give the pure active carbonate (26) (98 mg, 75%) as an orange oil. ^{1}H-NMR ($CDCl_3$, 300 MHz): δ 2.17 (s, 3H); 2.55 (s, 3H); 3.55 (s, 2H); 3.73 (s, 3H); 3.77 (m, 4H); 3.83 (m, 2H); 3.89 (t, *J* 5 Hz, 2H); 4.23 (t, *J* 5 Hz, 2H); 4.43 (m, 2H); 6.82 (d, *J* 8 Hz, 1H); 7.10 (s, 1H); 7.19 (d, *J* 12 Hz, 1H); 7.29 (d, *J* 8 Hz, 2H); 7.45 (d, *J* 9.5 Hz, 2H); 7.44 (d, *J* 8 Hz, 2H); 8.24 (d, *J* 9.5 Hz, 2H).

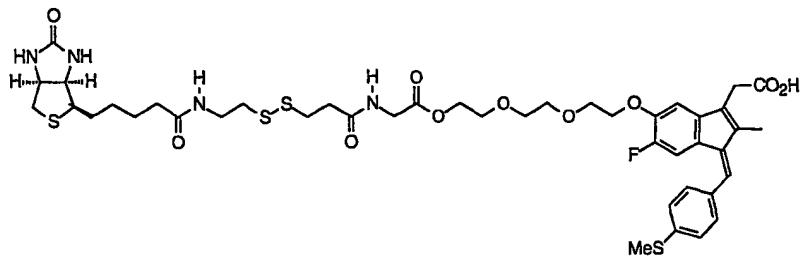
Example 1.17 – Synthesis of TentaGel bound methyl 2-(6-fluoro-5-{2-[2-(2-hydroxyethoxy)ethoxy]ethoxy}-2-methyl-1-[(4-methylthiophenyl)methylene]inden-3-yl)acetate (27)

TentaGel S-NH_2 (Fluka, *ca.* 0.45 mmol N/g resin, particle size 150-200 mm; 50 mg) was added to a solution of the carbonate (26) (20 mg, 30 mmol) and *N*-methylmorpholine (30 mg, 0.30 mmol) in dimethylformamide (1 cm^3). The mixture was shaken at room temperature for 16 h and filtered. The resin was washed with dimethylformamide (3x) followed by methanol (3x) then dried *in vacuo* to give 61 mg of the resin (27). IR (potassium bromide disc): ν 1734, 1718 cm^{-1}.

Example 1.18 – Synthesis of TentaGel bound 2-(6-fluoro-5-{2-[2-(2-hydroxyethoxy)ethoxy]ethoxy}-2-methyl-1-[(4-methylthiophenyl)methylene]inden-3-yl)acetic Acid (28)

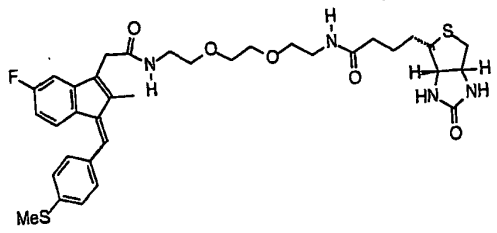
The resin (27) (58 mg) was added to a 0.25 M solution of sodium hydroxide in 2:1 ethanol water (4 cm^3). The mixture was shaken at room temperature for 4 h and filtered. The resin was washed with water (3x) then 1 M hydrochloric acid (3x), water (3x) and finally methanol (3x). Drying *in vacuo* gave 58 mg of the modified resin. IR (potassium bromide disc): ν 1718 cm^{-1}.

Alternatively, the sulindac analogue (IV) was conjugated with glycine and the amine thus produced was coupled to a commercially available biotin derivative to give the biotin labelled compound (29). The biotin derivative was then attached to a solid support by way of the biotin moiety using a biotin:streptavidin coupling technique.



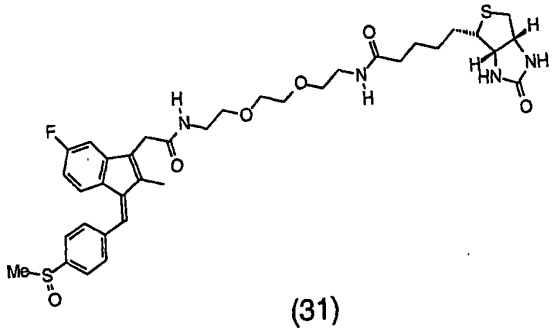
(29)

The biotinylated derivative (30) of sulindac sulfide was also produced by carbodiimide mediated coupling of sulindac sulfide with a commercially available biotin derivative.

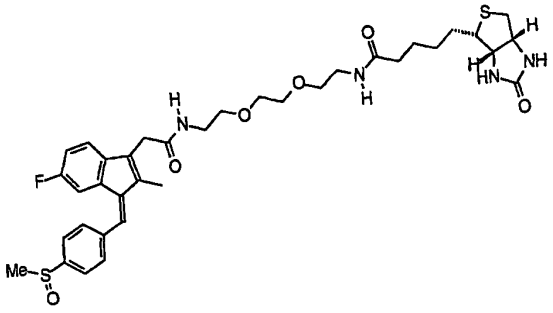


(30)

Further biotinylated derivatives, (31) and (32), were produced *via* the same chemistry and may be coupled to a solid phase by way of biotin:streptavidin coupling.



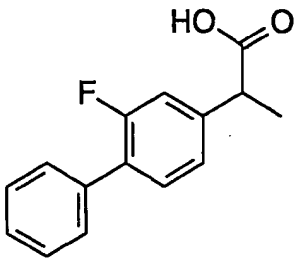
(31)



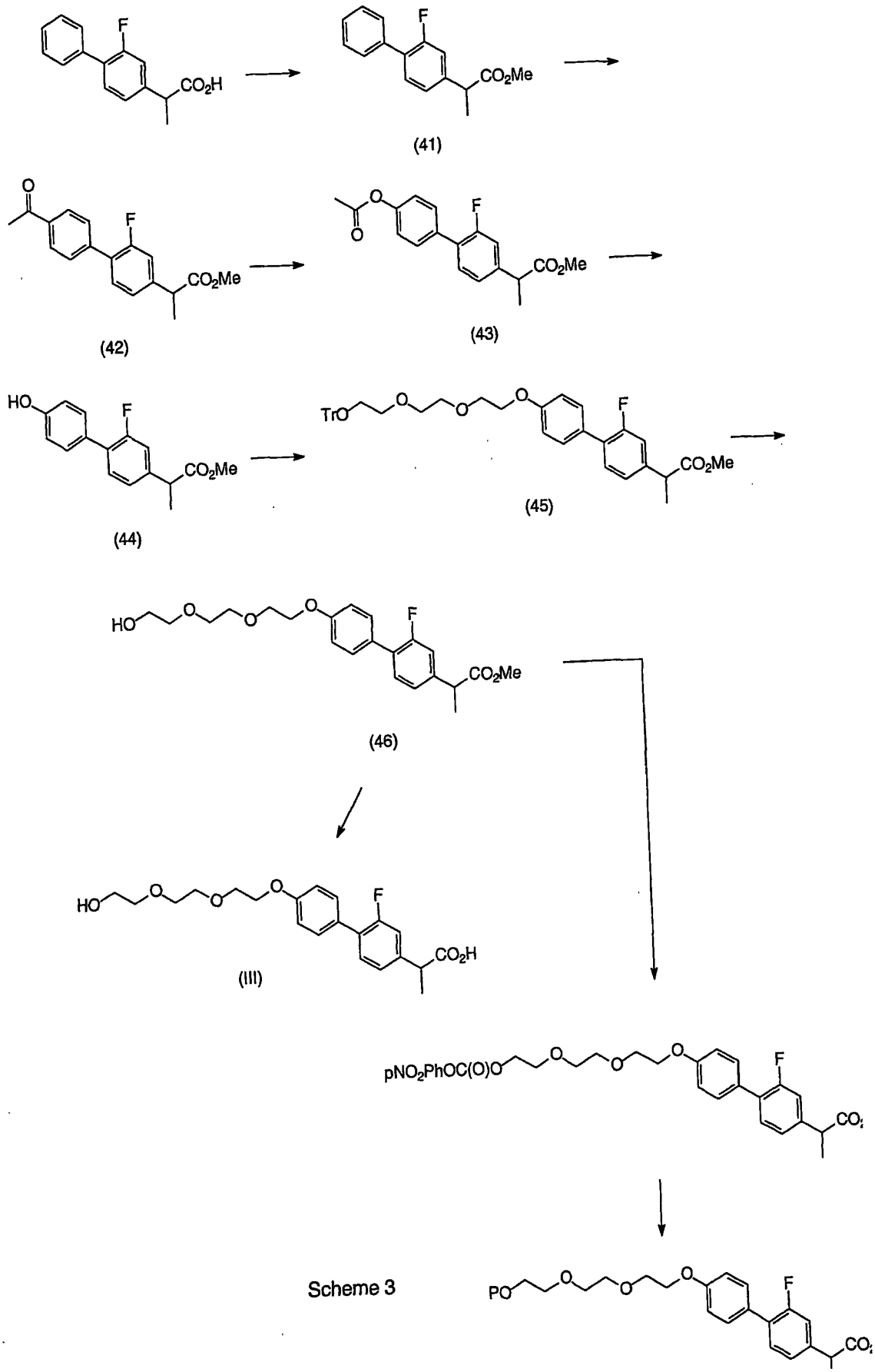
(32)

Example 2 - Synthesis of analogues of flurbiprofen and attachment to a solid phase

Flurbiprofen is a drug that acts to decrease the number of precancerous lesions (adenomas) in the colon in animals. The structure of flurbiprofen is as follows:



Compounds of formula (III) were produced by following the method provided in Scheme 3.



Scheme 3

Example 2.1 – Synthesis of methyl (2R,S)-(3-fluoro-4-phenylphenyl)propanoate (41)

Methanolic hydrogen chloride was generated by the careful addition of thionyl chloride (0.57 g, 4.7 mmol) to methanol (16 cm^3) at 5°. (2*R,S*)-(3-fluoro-4-phenylphenyl)propanoic acid (Flurbiprofen; 1.00 g, 4.09 mmol) was added and the solution stirred at 5° for 4 h, followed by stirring for a further 15 h at room temperature. The methanol was removed *in vacuo* to give the methyl ester (41) (1.06 g, quantitative) as a colourless oil. ^{1}H-NMR ($CDCl_3$, 200 MHz): δ 1.54 (d, *J* 7 Hz, 3H); 3.69 (s, 3H); 3.75 (q, *J* 7 Hz, 1H); 7.0-7.6 (m, 8H).

Example 2.2 – Synthesis of methyl (2R,S)-[4-(4-acetylphenyl)-3-fluorophenyl]propanoate (42)

A solution of methyl ester (41) (1.00 g, 3.87 mmol) and acetyl chloride (0.58 g, 7.4 mmol) in dry dichloromethane (5 cm^3) was added slowly to a stirred slurry of aluminium chloride (1.40 g, 10.5 mmol) and dry dichloromethane (5 cm^3) at 0°. The mixture was stirred at room temperature for 2 h and then poured on to ice (25 g). The phases were separated and the aqueous phase was washed twice with more dichloromethane. All of the organic washings were combined and dried ($MgSO_4$). Filtration and removal of the solvents *in vacuo* gave the ketone (42) (0.976 g, 84%) as a yellow oil. ^{1}H-NMR ($CDCl_3$, 200 MHz): δ 1.52 (d, *J* 7 Hz, 3H); 2.62 (s, 3H); 3.69 (s, 3H); 3.76 (q, *J* 7 Hz, 1H); 7.10 (m, 2H); 7.40 (t, *J* 8 Hz, 1H); 7.61 (m, 2H); 8.00 (m, 2H).

Example 2.3 – Synthesis of methyl (2R,S)-[4-(4-acetyloxyphenyl)-3-fluorophenyl]propanoate (43)

Meta-Chloroperoxybenzoic acid (Aldrich 77%; 1.495 g contained a maximum of 1.151 g, 6.67 mmol of peracid) was added in one portion to a stirred solution of

the ketone (42) (1.00 g, 3.33 mmol) in dry dichloromethane (7 cm^3). The mixture was cooled in an ice bath and trifluoroacetic acid (379 mg, 3.32 mmol) was added over a 20 min period. Removal of the ice bath was followed by stirring at room temperature for 3 days, at which time t.l.c. analysis indicated that consumption of starting material was complete. The mixture was diluted with more dichloromethane (20 cm^3) and washed with 10% aqueous sodium sulphite (10 cm^3), followed by saturated aqueous sodium carbonate (10 cm^3) and water (2x10 cm^3). The dichloromethane solution was dried ($MgSO_4$), filtered and evaporated *in vacuo* to give the diester (43) (754 mg, 72%) as a white solid. ^{1}H-NMR ($CDCl_3$, 200 MHz): δ 1.59 (d, *J* 7 Hz, 3H); 2.39 (s, 3H); 3.76 (s, 3H); 3.82 (q, *J* 7 Hz, 1H); 7.20 (m, 4H); 7.28 (t, *J* 8 Hz, 1H); 7.60 (m, 2H).

Example 2.4 – Synthesis of methyl (2R,S)-[3-fluoro-4-(4-hydroxyphenyl)phenyl]propanoate (44)

Water (30 cm^3) and saturated aqueous sodium bicarbonate (23 cm^3) were added to a solution of the phenolic acetate (43) (1.00 g, 3.16 mmol) in ethanol (100 cm^3). The mixture was stirred at room temperature for 15 h then its pH was adjusted to 4 with 10% hydrochloric acid. Most of the ethanol was removed *in vacuo* and the residue extracted with ethyl acetate (3x200 cm^3). The combined organic extracts were dried ($MgSO_4$), filtered and the solvent evaporated *in vacuo*. The crude product was purified by Flash Chromatography (30% ethyl acetate/70% hexane) to give the pure phenol (44) (613 mg, 71%) as pale yellow crystalline solid. ^{1}H-NMR (Acetone-d_6, 200 MHz): δ 1.47 (d, *J* 7 Hz, 3H); 3.64 (s, 3H); 3.83 (q, *J* 7 Hz, 1H); 6.91 (m, 2H); 7.16 (m, 2H); 7.41 (m, 3H); 8.53 (s, 1H).

Example 2.5 – Synthesis of methyl (2R,S)-{3-fluoro-4-[4-(2-{2-[2-(triphenylmethoxy)ethoxy]ethoxy}ethoxy)phenyl]phenyl}propanoate (45)

Diethyl azadicarboxylate (404 mg, 2.32 mmol) was added slowly to a stirred solution of the alcohol (41) (909 mg, 2.32 mmol), the phenol (44) (600 mg, 2.19

mmol) and triphenylphosphine (608 mg, 2.32 mmol) in dry tetrahydrofuran (30 cm^3) at 0°. Stirring was continued at 0° for 3 h and then for a further 16 h at room temperature. The tetrahydrofuran was removed *in vacuo* and the residue taken up into a minimal amount of 2:3 ethyl acetate/hexane. Filtration removed the resultant white precipitate and the dissolved crude product was purified by Flash Chromatography (2:3 ethyl acetate/hexane) to give the pure phenolic ether (45) (895 mg, 63%) as pale yellow oil. ^{1}H-NMR ($CDCl_3$, 200 MHz): δ 1.51 (d, *J* 7 Hz, 3H); 3.23 (t, *J* 5 Hz, 2H); 3.67 (s, 3H); 3.72 (m, 7H); 3.88 (t, *J* 5 Hz, 2H); 4.14 (t, *J* 5 Hz, 2H); 6.93 (m, 2H); 7.10 (m, 2H); 7.30 (m, 10H); 7.44 (m, 8H).

Example 2.6 – Synthesis of methyl (2R,S)-[3-fluoro-4-(4-{2-[2-(2-hydroxyethoxy)ethoxy]ethoxy}phenyl)phenyl]propanoate (46)

A solution of the trityl ether (45) (6.94 g, 10.7 mmol) in ether (21 cm^3) and formic acid (21 cm^3) was stirred at room temperature for 7 min. The mixture was diluted with ether (100 cm^3) and washed with saturated aqueous sodium chloride solution. Washing with saturated aqueous sodium bicarbonate solution was repeated until carbon dioxide was no longer evolved. After washing again with saturated aqueous sodium chloride solution, the ether solution was dried ($MgSO_4$) and filtered. Evaporation of the ether *in vacuo* gave a crude product which was purified by Flash Chromatography (ethyl acetate) to give the pure alcohol (46) (3.14 g, 72%) as a colourless oil. ^{1}H-NMR ($CDCl_3$, 200 MHz): δ 1.51 (d, *J* 7 Hz, 3H); 3.61 (m, 2H); 3.68 (s, 3H); 3.71 (m, 7H); 3.87 (t, *J* 5 Hz, 2H); 4.17 (t, *J* 5 Hz, 2H); 6.97 (m, 2H); 7.09 (m, 2H); 7.34 (t, *J* 8 Hz, 1H); 7.44 (m, 2H).

Example 2.7 – Synthesis of (2R,S)-[3-fluoro-4-(4-{2-[2-(2-hydroxyethoxy)ethoxy]ethoxy}phenyl)phenyl]propanoic acid (III)

The methyl ester (46) (1.60 g, 3.94 mmol) was dissolved in a solution of sodium hydroxide (520 mg, 13.0 mmol) in water (6.5 cm^3) and ethanol (13 cm^3). The solution was stirred at room temperature for 2 h and then adjusted to pH 7 with

hydrochloric acid. The ethanol was removed *in vacuo* and the residue extracted with ether. The ether solution was dried ($MgSO_4$), filtered and the solvent removed to yield the acid (1.05 g, 68%) as an off-white solid. The aqueous layer was acidified to pH 1 and again extracted with ether; processing of the ether solution as above yielded more of the acid (III) (389 mg, 25%) as a white solid. Both samples of the acid gave identical ^{1}H-NMR spectra. ^{1}H-NMR (Acetone-d_6, 200 MHz): d 1.49 (d, *J* 7 Hz, 3H); 3.56 (m, 2H); 3.67 (m, 7H); 3.84 (t, *J* 5 Hz, 2H); 4.18 (t, *J* 5 Hz, 2H); 7.16 (m, 2H); 7.23 (m, 2H); 7.44 (t, *J* 8 Hz, 1H); 7.51 (m, 2H).

Attachment of compound (III) to the solid phase was achieved by conversion to the corresponding 4-nitrophenyl carbonate, followed by reaction with TentaGel S-NH_2 to give compound (50).

Example 2.8 – Synthesis of methyl (2R,S)-{3-fluoro-4-[4-(2-{2-[2-(4-nitrophenoxycarbonyloxy)ethoxy]ethoxy}ethoxy)phenyl]phenyl}propanoate (48)

4-Nitrophenyl chloroformate (209 mg, 1.04 mmol) was added in one portion to a stirred solution of the alcohol (III) (211 mg, 0.519 mmol), *N*-methylmorpholine (158 mg, 1.56 mmol) and dimethylaminopyridine (6.3 mg, 52 mmol) in dichloromethane (5 cm^3) at 0°. The ice bath was removed and the mixture was stirred for 15 h. The dichloromethane was removed *in vacuo* and ether and 1 M hydrochloric acid were added. The layers were separated and the ether phase washed with more 1 M hydrochloric acid, followed by saturated aqueous sodium hydrogen carbonate and finally with saturated aqueous sodium chloride. Drying ($MgSO_4$), filtration and removal of solvent *in vacuo* gave the crude product which was purified by radial chromatography ("Chromatotron") using an ethyl acetate/hexane gradient to give the pure active carbonate (48) (192 mg, 65%) as a colourless oil. ^{1}H-NMR ($CDCl_3$, 200 MHz): δ 1.53 (d, *J* 7 Hz, 3H); 3.70 (s, 3H); 3.76 (m, 5H); 3.82 (m, 2H); 3.90 (m, 2H); 4.17 (m, 2H); 4.45 (m, 2H); 6.97 (m, 2H); 7.09 (m, 2H); 7.3-7.5 (m, 3H); 7.37 (d, *J* 9 Hz, 2H); 8.26 (d, *J* 9 Hz, 2H).

Example 2.9 – Synthesis of TentaGel bound methyl (2R,S)-[3-fluoro-4-(4-{2-[2-(2-hydroxyethoxy)ethoxy]ethoxy}phenyl)phenyl]propanoate (49)

TentaGel S-NH_2 (Fluka, *ca.* 0.45 mmol N/g resin, particle size 150-200 mm; 106 mg) was added to a solution of the carbonate (48) (68 mg, 120 mmol) and *N*-methylmorpholine (100 mg, 0.99 mmol) in dimethylformamide (5 cm^3). The mixture was shaken at room temperature for 16 h and filtered. The resin was washed with a 5% (v/v) solution of diisopropylethylamine in dimethylformamide (5x) followed by dimethylformamide (3x) and methanol (3x), then dried *in vacuo* to give 114 mg of the modified resin (49). IR (potassium bromide disc): ν 1734, 1718 cm^{-1}.

Example 2.10 – Synthesis of TentaGel bound (2R,S)-[3-fluoro-4-(4-{2-[2-(2-hydroxyethoxy)ethoxy]ethoxy}phenyl)phenyl]propanoic acid (50)

The resin (49) (80 mg) was added to a 0.50 M solution of sodium hydroxide in 2:1 ethanol water (3 cm^3). The mixture was shaken at room temperature for 4 h and filtered. The resin was washed with water (3x) then 1 M hydrochloric acid (3x), water (3x) and finally methanol (3x). Drying *in vacuo* gave 78 mg of the resin (50). IR (potassium bromide disc): ν 1718 cm^{-1}.

Example 3 - Phage display library construction

Phage display libraries were constructed using standard protocols for directional cloning of cDNA. The bacteriophage phage display lambda vectors were prepared as outlined below. The vectors used were T7 Select 1-1b, T7 Select 10-3, λfooDc and λvsx.1. cDNA libraries were made using RNA isolated from human adenomas and normal colonic tissue. The use of two different cDNA libraries not only allows the identification of proteins in a particular sample that bind to a target molecule in the presence of a second molecule, but also allows the identification of any binding proteins that may be differentially represented between the two populations of cells.

cDNA for cloning into the T7 Select vector was synthesised using standard procedures with *Hind*III random primers and *EcoR*I/*Hind*III linkers. cDNA for cloning into λfooDc was synthesised using *EcoR*I random primers (5' TCNNNNNN 3') and *Hind*III/*EcoR*I linkers (5' ATTCAAGCTTGAAT 3').

cDNA for cloning into λvsx.1 was synthesised using *Not*I random primers (5' GCNNNNNN 3') and *EcoR*I/*Not*I linkers (5' GGCCGCGAATTCGCGGCC 3').

Size selection of cDNA was performed on Size-Sep 400 columns according to standard protocols.

Vector arms and cDNA were ligated overnight at 16°C, at a cDNA to arms ratio of 0.5μg:25ng in a volume of 10μl or less. Ligations were packaged according to standard protocols. The libraries were titred, amplified and stored using standard protocols.

The preparation of vectors λvsx.1 and λfooDc were as follows: bacteriophage DNA was isolated by the plate lysate method, by plating 20 x 150mm dishes for confluent lysis (4 x 10^5pfu/dish) using LE392MP as host cells. The phage was eluted in SM buffer (10ml/dish) for 6 hours, harvested and 1% v/v $CHCl_3$ added. The eluate was spun at 3000rpm for 10 minutes and the supernatant recovered. The DNase I concentration was adjusted to 2 μg/ml and incubated at 37°C for 60 minutes. The mixture was adjusted to 10% w/v PEG8000, 100mM NaCl by dissolution and phage precipitated at 4°C overnight. The mixture was spun at 4500rpm for 20 minutes and the phage pellet resuspended in 4ml SM buffer and transfered to 1.5ml microfuge tubes. A spin of 2 minutes at room temperature was performed, supernatant pooled and TEAE-cellulose pre-equilibrated as follows was added:

Wash 1.5 g dry weight resin in 3 x 10 ml SM buffer

Adjust to 80% slurry

Rotate for 10 minutes

Spin at 3000rpm for 10 minutes and filter supernatant (0.45 μm).

Proteinase K and EDTA were added to a final concentration of 50µg/ml and 20mM respectively, incubated at 45°C for 30 minutes and 0.05 volumes 5%CTAB/0.5M NaCl added. This was incubated at 68°C for 10 minutes, chilled on ice for 2 minutes, spun in a microfuge at room temperature for 10 minutes the pellet resuspended in 1.2M NaCl and add 2.5 volumes EtOH added, spun 10 minutes at room temperature and the pellet resuspended in TE pH 8.0. The mixture was extracted with buffered phenol/ $CHCl_3$ and the aqueous phase recovered. DNA was precipitated with 0.1 volumes 3M NaOAc and 2.5 volumes EtOH, resuspended in TE pH 8.0 and concentration adjusted to 0.5 µg/µl

The *cos* ends were ligated as follows:

Mix 42µl λ DNA (0.5µg/µl)
5µl 10 x Ligation Buffer (including DTT and ATP)
10 units T4 DNA Ligase
H_2O to 50µl

The reaction was incubates at room temperature overnight (reaction becomes very viscous)

Vector Arms were prepared as follows:

λvsx.1 Digest with *EcoR*I and *Not*I
λfooDc Digest with *Hind*III and *EcoR*I

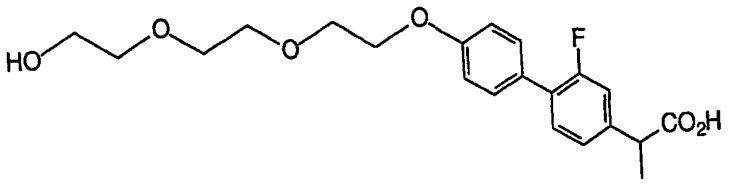
Ligation reactions were adjusted to 1 x *EcoR*I buffer in a final volume of 200 µl and included 40 units each of the appropriate enzymes (also include 0.1mg/ml BSA in the λvsx.1 digest). Reactions were incubated at 37°C overnight. A further 10 units of each enzyme was added and incubated overnight at 37°C.

Vectors were then treated with alkaline phosphatase to reduce background of non-recombinants in the library. Ligation/digestion was monitored by agarose gel electrophoresis. The vector was extracted with buffered phenol/ $CHCl_3$ and the aqueous phase recovered. DNA was precipitated with 0.1 volumes 3M NaOAc

and 2.5 volumes EtOH and resuspend in TE pH8.0. Concentration was adjusted to 0.5µg/µl.

Example 4 - Library screening for proteins that bind to a specific drug coupled to a solid support, in the presence of a free drug analogue

Binding of proteins displayed on the surface of a phage to the (R)-isomer of the following molecule in the presence of the inactive (S)-isomer:



Resin alone with linker attached, and drug covalenty coupled to coated resin, were blocked with 2% skim milk in binding buffer by constant mixing for 1 hour at room temperature. The resins were washed 5 times with 1.5 ml of binding buffer (20 mM Tris HCl pH 7.5, 0.25 M NaCl, 0.1% Tween 20).

10^{12} phage particles (T7 Select 1-1b, T7 Select 10-3, λfooDc or λvsx.1) containing cloned adenoma and normal colon cDNA libraries were incubated with resin (25 mg) plus linker in binding buffer with constant mixing for 1 hour at room temperature and the phage containing supernatant recovered, in order to remove any phage that may bind to the resin/linker alone.

The resultant phage supernatant was then incubated with drug-coated resin (20 mg) at room temperature with constant mixing for 2 hours in the presence of 0.5 to 5 mM inactive drug analogue, being the (S)-isomer of the above molecule.

The drug-coated resin was washed 5 times with 1.5 ml of binding buffer. The bound phage were eluted from the resin by washes with increasing concentrations of free drug from 10 nM-100 µM, using ten fold increments in binding buffer over a total period of 2-16 hours at room temperature with

constant mixing. For T7 phage a further elution of the resin with 1% SDS for 5 minutes at room temperature was required.

For re-use of the resin for subsequent steps, the resin is stripped with 5% SDS, washed 5 times with binding buffer, and stored at 4°C for next round of selection.

Eluted phage were titred, amplified by re-infection of a competent host and the newly enriched pool phage titred using standard protocols.

Twenty plaques so isolated were picked and amplified by PCR using synthetic oligonucleotide primers that flank the vector polylinker site. The size of inserts was determined by agarose gel electrophoresis and the sequence of inserts was determined by standard procedures using an Applied Biosystems 310 automatic DNA sequencing machine.

Cycles of the above process of binding phage to the solid support with coupled drug, in the presence of the analogue, were reiterated until multiple representations of the same insert were present in the purified phage pool. Typically 5-10 rounds of selection were required.

The oligonucleotides used for amplifying inserts from T7 vectors are commercially available. The specific oligonucleotides used for amplifying inserts from λfooDc are:

5'-GACCGTTGGGCCAATTGTC and 5'-TAAAACGACGGCCAGTGCC

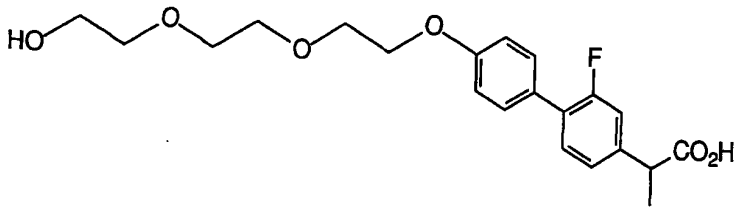
Oligonucleotides used for amplifying inserts from λvsx.1 are:

5'-AAATTACCGTCACCGCCAGT and 5'-TTTGATGCCTGGCAGTTCC

The aim of this screening strategy is to not only identify within the one experiment phage that may bind to the drug, but also to identify phage that may bind with differing affinities.

Example 6 - Library screening for proteins that bind to a specific biotinylated drug in the presence of a drug analogue

Binding of proteins displayed on the surface of a phage to the (R)-isomer of the following drug molecule in the presence of the inactive (S)-isomer:



10^{12} phage particles (T7 Select 1-1b, T7 Select 10-3, λfooDc or λvsx.1) with the adenoma and normal colon cloned cDNA libraries inserted were bound to a biotinylated version of the above drug in the range of concentrations from 10 nM – 50 μM and in the presence of 0.5 – 5 mM of the (S)-isomer analogue, and the complexes captured with 0.25 ml of paramagnetic streptavidin particles (Dynal).

The virus particles were washed 5 times with binding buffer (5x volume) at room temperature and phage eluted with 200 μM drug in binding buffer for 2-16 hours at room temperature. Alternatively, the streptavidin particles may be saturated with biotinylated drug, 0.45 ml of 2 μM drug per 0.1 ml of particles.

The bound phage were then eluted from the paramagnetic particles by washes with increasing concentrations of free drug in binding buffer from 10 nM-100 μM, using ten fold increments over a total period of 2-16 hours at room temperature with constant mixing.

Eluted phage were titred, amplified by re-infection of a competent host and the new enriched pool phage titred using standard protocols.

Twenty plaques so isolated were picked and amplified by PCR using synthetic oligonucleotide primers that flank the vector polylinker site. The size of inserts was determined by agarose gel electrophoresis and the sequence of inserts was determined by standard procedures using an Applied Biosystems 310 automatic DNA sequencing machine.

Cycles of the above process of binding phage to the solid support with coupled drug, in the presence of the analogue, were reiterated drug until multiple representations of the same insert were present in the purified phage pool. Typically 5-10 rounds of selection were required.

The oligonucleotides used for amplifying inserts from T7 vectors are commercially available. For the other vectors, the following oligonucleotides are suitable:

The specific oligonucleotides used for amplifying inserts from λfooDc are:
5'-GACCGTTGGGCCAATTGTC and 5'-TAAAACGACGGCCAGTGCC
Oligonucleotides used for amplifying inserts from λvsx.1 are:
5'-AAATTACCGTCACCGCCAGT and 5'-TTTGATGCCTGGCAGTTCC

The aim of this screening strategy was to not only identify within the one experiment phage that may bind to the drug, but also to identify phage that may bind with differing affinities.

Finally, it is to be understood that various other modifications and/or alterations may be made without departing from the spirit of the present invention as outlined herein.

Claims:

1. A method for identifying a protein capable of binding to a target molecule, the method including the steps of:

(a) providing a pool of candidate proteins;

(b) providing a non-nucleic acid target molecule, wherein the non-nucleic acid target molecule is coupled to a selectable moiety;

(c) providing a second molecule which is structurally similar to the non-nucleic acid target molecule, wherein the second molecule is deficient in a desired activity of the target molecule;

(d) allowing one or more of the candidate proteins to bind to the non-nucleic acid target molecule in the presence of the second molecule;

(e) isolating a protein bound to the target molecule; and

(f) identifying the binding protein.

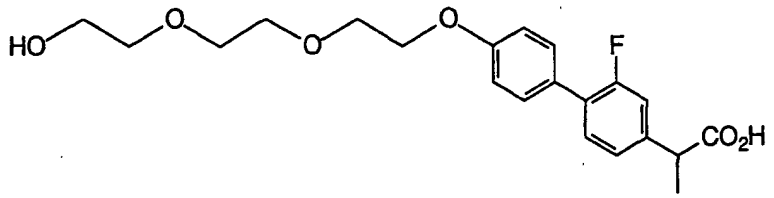
2. A method according to claim 1, wherein the non-nucleic acid target molecule is selected from the group consisting of drug molecules, proteins, peptides, polysaccharides, glycoproteins, hormones, receptors, lipids, small molecules, metabolites, cofactors, transition state analogues and toxins.

3. A method according to claim 2, wherein the non-nucleic acid target molecule is a drug molecule.

4. A method according to any one of claims 1 to 3, wherein the second molecule is an isomer, geometric isomer, enantiomer, conformer, stereoisomer, structural isomer or a chemically substituted derivative of the non-nucleic acid target molecule.

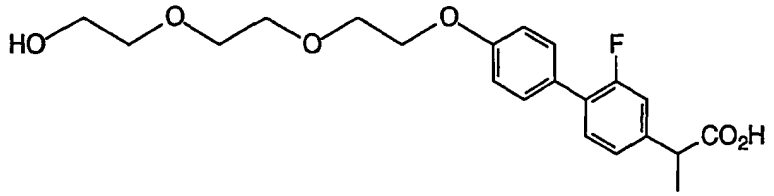
5. A method according to claim 3, wherein the drug molecule is flurbiprofen.

6. A method according to claim 3, wherein the target molecule is the (*R*)-stereoisomer of a molecule with the following formula:



7. A method according to claim 5, wherein the second molecule is an isomer, geometric isomer, enantiomer, conformer, stereoisomer, structural isomer or a chemically substituted derivative of flurbiprofen.

8. A method according to claim 6, wherein the second molecule is the (*S*)-stereoisomer of a molecule with the following formula:



9. A method according to claim 3, wherein the drug molecule is suldinac sulfide.

10. A method according to claim 9, wherein the second molecule is a isomer, geometric isomer, enantiomer, conformer, stereoisomer, structural isomer or a chemically substituted derivative of sulindac sulfide.

11. A method according to any one of claims 1 to 10, wherein the second molecule is present in a molar excess to the target molecule.

12. A method according to claim 11, wherein the second molecule is present in a molar excess of one hundred fold or greater relative to the target molecule.

13. A method according to any one of claims 1 to 12, wherein the pool of candidate proteins is expressed from DNA molecules inserted into a viral genome.

14. A method according to claim 13, wherein each of the candidate proteins is displayed on the surface of a viral particle.

15. A method according to claim 14, wherein the viral particle is derived from a bacteriophage.

16. A method according to claim 15, wherein the bacteriophage is selected from the group consisting of T7, T4, lambda, lambdoid phage, or filamentous phage.

17. A method according to any one of claims 1 to 16, wherein the selectable moiety is a biotin containing group or an activated carbonate group.

18. A method for identifying a protein capable of binding to a target molecule, the method including the steps of:

(a) providing a pool of candidate proteins, wherein each candidate protein is displayed on the surface of a viral particle;

(b) providing a non-nucleic acid target molecule, wherein the non-nucleic acid target molecule is coupled to a selectable moiety;

(c) providing a second molecule which is structurally similar to the non-nucleic acid target molecule, wherein the second molecule is deficient in a desired activity of the target molecule;

(d) allowing one or more of the candidate proteins to bind to the non-nucleic acid target molecule in the presence of the second molecule;

(e) isolating one or more proteins bound to the target molecule;

(f) amplifying the viral particles encoding the isolated binding proteins;

(g) reiterating steps (a) through (f); and

(h) identifying the binding protein.

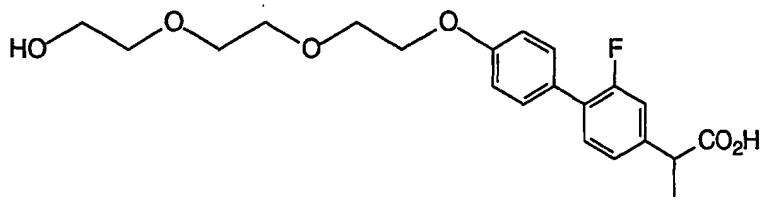
19. A method according to claim 18, wherein the non-nucleic acid target molecule is selected from the group consisting of drug molecules, proteins, peptides, polysaccharides, glycoproteins, hormones, receptors, lipids, small molecules, metabolites, cofactors, transition state analogues and toxins.

20. A method according to claim 19, wherein the non-nucleic acid target molecule is a drug molecule.

21. A method according to any one of claims 18 to 20, wherein the second molecule is an isomer, geometric isomer, enantiomer, conformer, stereoisomer, structural isomer or a chemically substituted derivative of the non-nucleic acid target molecule.

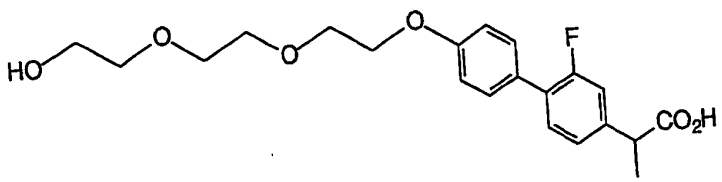
22. A method according to claim 20, wherein the drug molecule is flurbiprofen.

23. A method according to claim 20, wherein the target molecule is the (*R*)-stereoisomer of a molecule with the following formula:



24. A method according to claim 20, wherein the second molecule is an isomer, geometric isomer, enantiomer, conformer, stereoisomer, structural isomer or a chemically substituted derivative of flurbiprofen.

25. A method according to claim 23, wherein the second molecule is the (*S*)-stereoisomer of a molecule with the following formula:



26. A method according to claim 20, wherein the drug molecule is suldinac sulfide.

27. A method according to claim 26, wherein the second molecule is a isomer, geometric isomer, enantiomer, conformer, stereoisomer, structural isomer or a chemically substituted derivative of sulindac sulfide.

28. A method according to any one of claims 18 to 27, wherein the second molecule is present in a molar excess to the target molecule.

29. A method according to claim 28, wherein the second molecule is present in a molar excess to the target molecule of one hundred fold or greater.

30. A method according to any one of claims 18 to 29, wherein the viral particle is derived from a bacteriophage.

31. A method according to claim 30, wherein the bacteriophage is selected from the group consisting of T7, T4, lambda, lambdoid phage, or filamentous phage.

32. A method according to any one of claims 18 to 31, wherein the selectable moiety is a biotin containing group or an activated carbonate group.

33. A method for identifying a protein capable of binding to target molecule, the method including the steps of:

(i) providing first and second pools of candidate proteins;

(j) providing a non-nucleic acid target molecule, wherein the non-nucleic acid target molecule is coupled to a selectable moiety;

(k) providing a second molecule which is structurally similar to the non-nucleic acid target molecule, wherein the second molecule is deficient in a desired activity of the target molecule;

(l) allowing one or more of the candidate proteins in the first pool to bind to the non-nucleic acid target molecule in the presence of the second molecule;

(m) isolating one or more proteins in the first pool that bind to the target molecule;

(n) allowing one or more of the candidate proteins in the second pool to bind to the non-nucleic acid target molecule in the presence of the second molecule;

(o) isolating one or more proteins in the second pool that bind to the target molecule; and

(p) comparing one or more proteins isolated from each of the first and second pools to identify a protein that is differentially represented between the first and second pools.

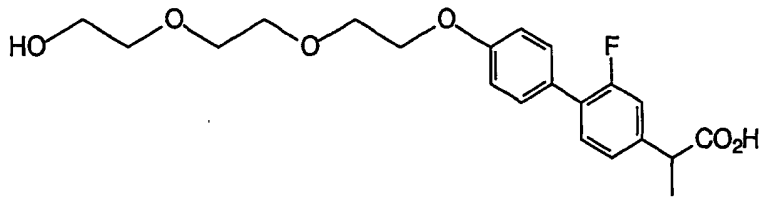
34. A method according to claim 33, wherein the non-nucleic acid target molecule is selected from the group consisting of drug molecules, proteins, peptides, polysaccharides, glycoproteins, hormones, receptors, lipids, small molecules, metabolites, cofactors, transition state analogues and toxins.

35. A method according to claim 34, wherein the non-nucleic acid target molecule is a drug molecule.

36. A method according to any one of claims 33 to 35, wherein the second molecule is an isomer, geometric isomer, enantiomer, conformer, stereoisomer, structural isomer or a chemically substituted derivative of the non-nucleic acid target molecule.

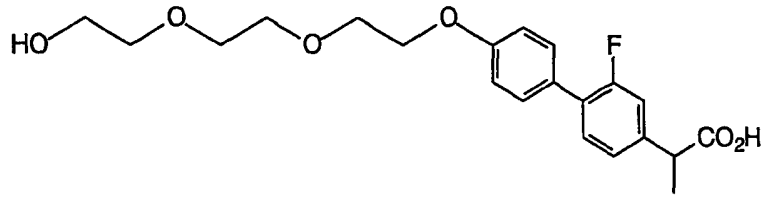
37. A method according to claim 35, wherein the drug molecule is flurbiprofen.

38. A method according to claim 35, wherein the target molecule is the (*R*)-stereoisomer of a molecule with the following formula:



39. A method according to claim 37, wherein the second molecule is an isomer, geometric isomer, enantiomer, conformer, stereoisomer, structural isomer or a chemically substituted derivative of flurbiprofen.

40. A method according to claim 38, wherein the second molecule is the (*S*)-stereoisomer of a molecule with the following formula:



41. A method according to claim 35, wherein the drug molecule is suldinac sulfide.

42. A method according to claim 41, wherein the second molecule is a isomer, geometric isomer, enantiomer, conformer, stereoisomer, structural isomer or a chemically substituted derivative of sulindac sulfide.

43. A method according to any one of claims 33 to 42, wherein the second molecule is present in a molar excess to the target molecule.

44. A method according to claim 43, wherein the second molecule is present in a molar excess to the target molecule of one hundred fold or greater.

45. A method according to any one of claims 33 to 45, wherein the pool of candidate proteins is expressed from DNA molecules inserted into a viral genome.

46. A method according to claim 45, wherein each of the candidate proteins is displayed on the surface of a viral particle.

47. A method according to claim 46, wherein the viral particle is derived from a bacteriophage.

48. A method according to claim 47, wherein the bacteriophage is selected from the group consisting of T7, T4, lambda, lambdoid phage, or filamentous phage.

49. A method according to any one of claims 33 to 48, wherein the selectable moiety is a biotin containing group or an activated carbonate group.

50. A method according to any one of claims 33 to 49, wherein the first pool of candidate proteins is derived from a cellular extract from non-cancerous cells and the second pool of candidate proteins is derived from a cellular extract from cancerous cells.

51. A method according to any one of claims 33 to 49, wherein the first pool of candidate proteins is derived from a cellular extract from non-cancerous cells and the second pool of candidate proteins is derived from a cellular extract from pre-cancerous cells.

52. A method according to any one of claims 33 to 49, wherein the first pool of candidate proteins is derived from a cellular extract from pre-cancerous cells and the second pool of candidate proteins is derived from a cellular extract from cancerous cells.

53. A method according to any one of claims 50 to 52, wherein the cellular extracts are derived from a human cell.

54. A method according to claim 53, wherein the human cell is derived from colorectal tissue, breast tissue, cervical tissue, uterine tissue, renal tissue, pancreatic tissue, oesophageal tissue, stomach tissue, lung tissue, brain tissue, liver tissue, bladder tissue, bone tissue, prostate tissue, skin tissue, ovary tissue, testicular tissue, muscle tissue or vascular tissue.

55. A method for identifying a protein capable of binding to target molecule, the method including the steps of:

(a) providing a first pool of candidate proteins;

(b) providing a non-nucleic acid target molecule, wherein the non-nucleic acid target molecule is coupled to a selectable moiety;

(c) providing a second molecule which is structurally similar to the non-nucleic acid target molecule, wherein the second molecule is deficient in a desired activity of the target molecule;

(d) allowing one or more of the candidate proteins in the first pool to bind to the non-nucleic acid target molecule in the presence of the second molecule;

(e) isolating a protein in the first pool that binds to the target molecule;

(f) comparing the level of the protein in the first pool of candidate proteins with the level of the protein in a second pool of proteins; and

(g) identifying a protein that is differentially represented between the first and second pools.

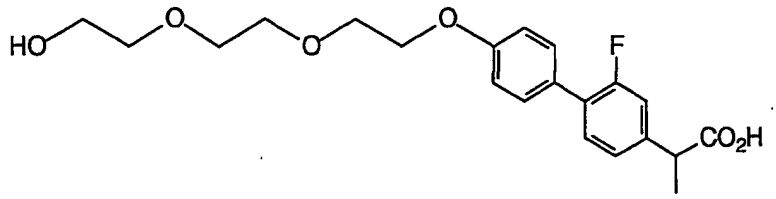
56. A method according to claim 55, wherein the non-nucleic acid target molecule is selected from the group consisting of drug molecules, proteins, peptides, polysaccharides, glycoproteins, hormones, receptors, lipids, small molecules, metabolites, cofactors, transition state analogues and toxins.

57. A method according to claim 56, wherein the non-nucleic acid target molecule is a drug molecule.

58. A method according to any one of claims 1 to 3, wherein the second molecule is an isomer, geometric isomer, enantiomer, conformer, stereoisomer, structural isomer or a chemically substituted derivative of the non-nucleic acid target molecule.

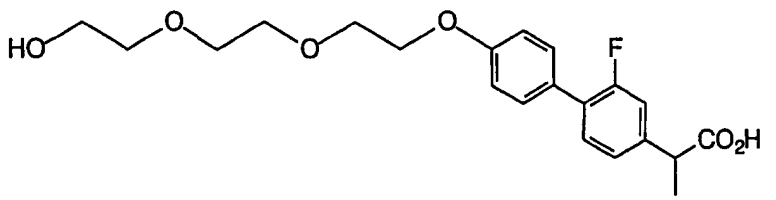
59. A method according to claim 57, wherein the drug molecule is flurbiprofen.

60. A method according to claim 57, wherein the target molecule is the (*R*)-stereoisomer of a molecule with the following formula:



61. A method according to claim 59, wherein the second molecule is an isomer, geometric isomer, enantiomer, conformer, stereoisomer, structural isomer or a chemically substituted derivative of flurbiprofen.

62. A method according to claim 60, wherein the second molecule is the (*S*)-stereoisomer of a molecule with the following formula:



63. A method according to claim 57, wherein the drug molecule is suldinac sulfide.

64. A method according to claim 63, wherein the second molecule is a isomer, geometric isomer, enantiomer, conformer, stereoisomer, structural isomer or a chemically substituted derivative of sulindac sulfide.

65. A method according to any one of claims 55 to 64, wherein the second molecule is present in a molar excess to the target molecule.

66. A method according to claim 65, wherein the second molecule is present in a molar excess to the target molecule of one hundred fold or greater.

67. A method according to any one of claims 55 to 66, wherein the pool of candidate proteins is expressed from DNA molecules inserted into a viral genome.

68. A method according to claim 67, wherein each of the candidate proteins is displayed on the surface of a viral particle.

69. A method according to claim 68, wherein the viral particle is derived from a bacteriophage.

70. A method according to claim 69, wherein the bacteriophage is selected from the group consisting of T7, T4, lambda, lambdoid phage, or filamentous phage.

71. A method according to any one of claims 55 to 70, wherein the selectable moiety is a biotin containing group or an activated carbonate group.

72. A method according to any one of claims 55 to 71, wherein the first pool of candidate proteins is derived from a cellular extract from non-cancerous cells and the second pool of candidate proteins is derived from a cellular extract from cancerous cells.

73. A method according to any one of claims 55 to 71, wherein the first pool of candidate proteins is derived from a cellular extract from non-cancerous cells and the second pool of candidate proteins is derived from a cellular extract from pre-cancerous cells.

74. A method according to any one of claims 55 to 71, wherein the first pool of candidate proteins is derived from a cellular extract from pre-cancerous cells and the second pool of candidate proteins is derived from a cellular extract from cancerous cells.

75. A method according to any one of claims 72 to 74, wherein the cellular extracts are derived from a human cell.

76. A method according to claim 75, wherein the human cell is derived from colorectal tissue, breast tissue, cervical tissue, uterine tissue, renal tissue, pancreatic tissue, oesophageal tissue, stomach tissue, lung tissue, brain tissue, liver tissue, bladder tissue, bone tissue, prostate tissue, skin tissue, ovary tissue, testicular tissue, muscle tissue or vascular tissue.

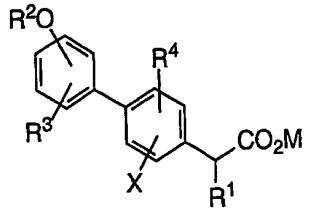
77. A protein identified according to the method of any one of claims 1 to 17.

78. A protein identified according to the method of any one of claims 18 to 32.

79. A protein identified according to the method of any one of claims 33 to 54.

80. A protein identified according to the method of any one of claims 55 to 76.

81. A compound with the following chemical formula:



(II)

or a salt thereof, wherein:

- R^1 is selected from hydrogen and lower alkyl (C1 to C8);
- R^2 is $YX^2((CH_2)_m X^2)_n$-, wherein m is 2 to 4, n is 1 to 6, X^2 is selected from O, S and N, and Y is independently selected from hydrogen, lower alkyl, or a suitable heteroatom protecting group;
- R^3 is selected from one or more of hydrogen, alkyl, aryl, halogen, hydroxy, alkoxy, aryloxy, amino (unsubstituted and substituted) and caroboxy;
- R^4 is selected from one or more of hydrogen, alkyl, aryl, halogen, hydroxy, alkoxy, aryloxy, amino (unsubstituted and substituted) and caroboxy;
- X is selected from fluoro, chloro, bromo and iodo;
- M is selected from hydroxy, alkoxy, aryloxy, amino, alkylamino (mono- and di-), arylamino (mono- and di-), N-morpholino, hydroxyalkylamino, dialkylaminoalkylamino, aminoalkylamino, polyhydroxyamino, and salts of any of the aforementioned.

82. A compound as in claim 81 wherein X is fluoro.

83. A compound as in claim 81 wherein the fluoro group is substituted *meta* to the alkylcarboxylate group.

84. A compound as in claim 81 wherein R^1 is a lower alkyl group.

85. A compound as in claim 84 wherein R^1 is a methyl group.

86. A compound as in claim 84 wherein R^2 is an alkyleneoxy or polyoxyalkylene chain.

87. A compound as in claim 86 wherein the alkyleneoxy or polyoxyalkylene chain has between 1 and 4 alkyleneoxy repeating units.

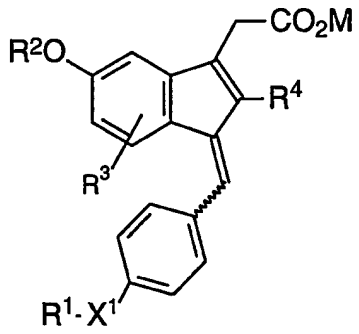
88. A compound as in claim 87 wherein R^2 is a triethylene glycol group.

89. A compound as in claim 86 wherein the R^2O- group is substituted at a position *para* to the aryl substituent.

90. A compound as in claim 89 wherein both R^3 and R^4 are hydrogen.

91. A compound as in claim 90 wherein M is hydroxy or a salt thereof.

92. A compound with the following chemical formula:



(II)

or a salt thereof, wherein:

- X^1 is selected from sulfide, sulfone and sulfoxide;
- R^1 is selected from hydrogen, hydroxy (when X^1 is sulfone or sulfoxide), and lower alkyl (C1 to C8);

- R^2 is $YX^2((CH_2)_m X^2)_n$-, wherein m is 2 to 4, n is 1 to 6, X^2 is selected from O, S and N, and Y is independently selected from hydrogen, lower alkyl, or a suitable heteroatom protecting group;
- R^3 is selected from hydrogen, halogen, alkyl, alkoxy, acyloxy, amino, alkylamino (mono- and di-), arylamino (mono- and di-), nitro, cyano, carboxyl;
- R^4 is selected from hydrogen and lower alkyl (C1 to C8); and
- M is selected from hydroxy, alkoxy, aryloxy, amino, alkylamino (mono- and di-), arylamino (mono- and di-), N-morpholino, hydroxyalkylamino, dialkylaminoalkylamino, aminoalkylamino, polyhydroxyamino, and salts of any of the aforementioned.

93. A compound as in claim 92 wherein X^1 is either a sulfone or a sulfide.

94. A compound as in claim 93 wherein X^1 is a sulfide.

95. A compound as in claim 94 wherein R^1 is a lower alkyl group.

96. A compound as in claim 95 wherein R^1 is a methyl group.

97. A compound as in claim 95 wherein R^2 is an alkyleneoxy or polyoxyalkylene chain.

98. A compound as in claim 97 wherein the alkyleneoxy or polyoxyalkylene chain has between 1 and 4 alkyleneoxy repeating units.

99. A compound as in claim 98 wherein R^2 is a triethylene glycol group.

100. A compound as in claim 97 wherein R^3 is a halogen group.

101. A compound as in claim 100 wherein R^3 is a fluoro group.

102. A compound as in claim 101 wherein the fluoro group is ortho to the hydroxy group.

103. A compound as in claim 100 wherein R^4 is a lower alkyl group.

104. A compound as in claim 103 wherein R^4 is methyl.

105. A compound as in claim 103 wherein M is hydroxy or a salt thereof.